The University Of Sheffield.

Access To Thesis.

A fully completed copy of this form must be submitted to Research & Innovation Services prior to the award of your degree. If you are submitting a hard copy of the thesis the form should be bound into the front of the thesis

SECTION 1: STUDENT DETAILS			
Family Name	McDool	First Name	Emily
Registration Number	130128326	Department	Economics
Thesis Title	Evaluation of neighbourhood, class setting and academy school effects on education outcome		

SECTION 2: THESIS SUBMISSION DETAILS – PLEASE SELECT ONE OF THE FOLLOWING OPTIONS	
☐	I am submitting in print format only for deposit in the University Library **(Note: this option only applies to students who initially registered prior to 2008)**
☑	I am submitting an eThesis only to the White Rose eTheses Online server. I confirm that the eThesis is a complete version of my thesis and no content has been removed
☐	I am submitting an eThesis to the White Rose eTheses Online server and also submitting in print format because I have removed some content from my eThesis

SECTION 3: EMBARGO DETAILS – PLEASE SELECT FROM THE FOLLOWING OPTIONS

Each Faculty has agreed a pre-approved embargo threshold (Arts & Humanities – 1 yr; Engineering – 1 yr; Medicine, Dentistry & Health – 2 yrs; Science – 5 yrs; Social Sciences – 3 yrs. Requests for embargoes that exceed the Faculty threshold will require Faculty approval. If you wish to request a longer embargo, please complete and submit the form available at: www.shef.ac.uk/ris/pgr/code/embargoes

Please note that if no boxes are ticked, you will have consented to your thesis being made available without any restrictions.

Should the thesis be embargoed? If 'Yes', please specify the length of embargo requested (in years)	Print Thesis	☑ No	☐ Yes	________ Years
	eThesis	☑ No	☐ Yes	________ Years

Reason for the embargo (please select from the following options):

☐	Third party copyright	☐	Commercial confidentiality
☐	Contains personal data	☐	Could prejudice national security
☐	Could endanger health and safety	☐	Exempt under another category listed in the FOI Act 2000
☐	Planned publication	☐	Other

SECTION 4: COPYRIGHT LICENCE OPTIONS – PLEASE SELECT ONE OF THE FOLLOWING

This thesis is protected by the Copyright Design and Patents Act 1988. No reproduction is permitted without consent of the author. It is recommended that you make your thesis available using a Creative Commons Licence http://creativecommons.org/about/license/. This Licence protects you as the author of the work and also clarifies the uses that others may make of your work.

☑	Creative Commons Attribution-Non-Commercial-No-derivatives **(recommended)**	☐	Creative Commons Attribution-Non-Commercial
☐	Creative Commons Attribution	☐	Creative Commons Attribution-No-derivative-Works
☐	Creative Commons Attribution-Non-Commercial-Share Alike	☐	Other/Do not apply a Licence

SECTION 5: THESIS DEPOSIT AGREEMENT - STUDENT		
1. I, the author, confirm that the Thesis is my own work, and that where materials owned by a third party have been used, copyright clearance has been obtained. I am aware of the University's *Guidance on the Use of Unfair Means* (www.sheffield.ac.uk/ssid/exams/plagiarism).		
2. I confirm that all copies of the Thesis submitted to the University, whether in print or electronic format, are identical in content and correspond with the version of the Thesis upon which the examiners based their recommendation for the award of the degree (unless edited as indicated above).		
3. I agree to the named Thesis being made available in accordance with the conditions specified above.		
4. I give permission to the University of Sheffield to reproduce the print Thesis (where applicable) in digital format, in whole or part, in order to supply single copies for the purpose of research or private study for a non-commercial purpose. I agree that a copy of the eThesis may be supplied to the British Library for inclusion on EThOS and WREO, if the thesis is not subject to an embargo, or if the embargo has been lifted or expired.		
5. I agree that the University of Sheffield's eThesis repository (currently WREO) will make my eThesis (where applicable) available over the internet via an entirely non-exclusive agreement and that, without changing content, WREO and/or the British Library may convert my eThesis to any medium or format for the purpose of future preservation and accessibility.		
6. I agree that the metadata relating to the eThesis (where applicable) will normally appear on both the University's eThesis server (WREO) and the British Library's EThOS service, even if the eThesis is subject to an embargo.		
Student's name (PLEASE PRINT): Emily McDool	Signature:	Date 30/6/16

SECTION 6: THESIS DEPOSIT AGREEMENT - SUPERVISOR		
I, the supervisor, agree to the named Thesis being made available in accordance with the conditions specified above.		
Supervisor's name (PLEASE PRINT): Steven McIntosh	Signature: S. McIntosh	Date: 4.7.16

SECTION 6: TO BE COMPLETED BY RESEARCH & INNOVATION SERVICES		
Does the embargo exceed the agreed Faculty length? FCA - 1YR; FCE - 1YR; FCM - 2YRS; FCP - 5 YRS; FCS - 3 YRS	☐ Yes* if 'yes' please attach embargo extension request form	☑ No
University stamp		

Evaluation of neighbourhood, class setting and academy school effects on education outcomes in the UK

By:

Emily Marie McDool

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Faculty of Social Science
Department of Economics

July 2016

ABSTRACT

This thesis includes three chapters that explore contemporary topics within the area of education in the UK.

The initial empirical chapter examines whether living in a deprived neighbourhood impacts upon the probability of obtaining the benchmark GCSE outcomes, when adopting a propensity score matching methodology. The chapter also examines whether there is a differential impact of neighbourhood deprivation upon children with educated parents, relative to those with uneducated parents. The results show that living in a deprived neighbourhood negatively influences the probability of gaining the observed GCSE outcomes; individuals with educated parents lose out to a greater extent by living in a deprived neighbourhood, relative to individuals with uneducated parents.

The subsequent chapter examines whether setting, which involves separating children into classes based on ability, influences the attitudes and behaviours of primary school children. A fixed effects methodology is initially adopted to identify the impact of being set in maths; the results signal that the behaviour of girls may be improved by setting. The chapter also investigates whether the level of the maths set in which the child is sorted influences behaviour by adopting an instrumental variables approach to overcome the likely endogeneity issue surrounding the set placement. The results indicate that whilst internalising behaviour was improved for girls placed in the lowest set, this set placement was detrimental to the internalising behaviour of boys.

The final chapter analyses the impact of post-2010 primary converter academies on pupil progress. Adopting a difference-in-difference methodology, individuals who experience academy conversion are compared with those whose school converted after leaving from the same school year cohort. The results indicate that converter academies had a positive impact upon pupil progress. When examining the effect by neighbourhood deprivation, the positive impact of converter academies is more consistent for schools in the least deprived neighbourhoods.

ACKNOWLEDGEMENTS

I would like to express my greatest sincere gratitude to my first supervisor Steve Mcintosh for his invaluable support, guidance and time throughout my PhD but also in the process of applying for the Masters and PhD programme so many years ago. Equally, I would like to thank my supervisor Gurleen Popli for her valued comments and guidance, but also for providing me with the opportunity to continue in my chosen career path. I also thank all other members of staff within the Department for Economics for their friendly support and encouragement, especially Jenny Roberts who gave me the chance to further my research experience. I would also like to show appreciation to the ESRC for financial support.

I am also grateful to all other PhD students, both past and present, especially Laura Requena, Damon Morris, Harry Pickard and James Oswald for providing advice and encouragement at all times. I would also like to express my gratitude to my supportive and dependable friends outside of academia, namely Katie Hanley, Sarah O'Brien, Fran Wilson, Amy Wilson and Danielle Keyworth.

Finally, and by no means least, I would like to express my appreciation to my family, especially my wonderful parents, Claire and Neal who have provided selfless and endless support, encouragement and inspiration at all times, and my sister Faye who often brings light and humour to my worries and concerns. I would finally like to give my greatest heartfelt thanks to my fiancé and best friend, John Hanrahan for his patience, understanding and complete support in every way possible.

NOTES AND DISCLAIMERS

The work of Chapter 2 was based on data from the Longitudinal Survey of Young People England (LSYPE), produced by the Department for Education (DfE) and supplied by the Secure Data Service at the UK Data Archive. The data are Crown Copyright and reproduced with the permission of the controller of HMSO and Queen's Printer for Scotland. The use of the data in this work does not imply the endorsement of ONS or the Secure Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 : INTRODUCTION

1.1 AIMS AND MOTIVATION

The economics of education literature has provided a greater insight into a multitude of topics surrounding education; early research in the field by Becker (1964) assisted in explaining why individuals choose to invest in education and training through the human capital theory, while Mincer (1958, 1974) developed the earnings function which has allowed educational economists to estimate the returns to education. More recently, there has been a growth in the economics of education research which has been attributed to improved data provision, policy relevance and development of methodological approaches (Machin, 2008). This surge in research has provided a greater understanding of issues such as intergenerational mobility and the impact of school choice and competition upon pupil outcomes, while policy evaluation research has become more relevant in informing policy makers.

In recent research, major developments have been made in explaining the determinants of both cognitive and non-cognitive pupil outcomes. Pupil outcomes, even at early stages in life, are likely to play a significant role in an individual’s life chances and adult outcomes; it is therefore imperative to understand the fundamental influences upon pupil outcomes, such as those associated with individual and family characteristics. Equally, it is important to understand how policies and interventions may be used to improve outcomes. Within the current economics of education literature, a number of determinants have been extensively examined, while a consensus has been reached regarding the impact of these determinants upon pupil outcomes. Progress has been made in explaining how factors such as family background influence individual outcomes, with research suggesting that children from deprived backgrounds and low income families generally complete school with considerably lower levels of educational attainment (Chowdry 2010; Hirsch 2007). It is also well established that gender and ethnicity influence pupil outcomes, as males attain lower test scores than females, while white children perform significantly worse at school than all other groups (Vignoles and Meschi, 2010). This is also true for summer born children who do worse than children born at the beginning of the academic year (Crawford 2010; Campbell 2013). Equally, the research suggests that birth order is a significant predictor of cognitive and non-cognitive outcomes (Sulloway 1996; Reinherz et al. 2003). In a similar manner, a number of school characteristics such as school resources and teacher quality have been examined as determinants of pupil outcomes (Hanushek et al., 2005) with factors such as class size providing the focus of debate within the existing literature (Hanushek, 2000).

Alongside individual and family characteristics, pupil outcomes are likely to be determined by factors which simultaneously influence other individuals and are associated with the pupils' environment outside of their household. Examples of such factors include neighbourhood characteristics, school-level policies and national-level educational policies. These topics form the focus of this thesis which will include three chapters that examine each of these factors in turn; specifically, the thesis will investigate the influence of neighbourhood deprivation, class setting by ability and converter academy schools upon a number of pupil outcomes. At present, the research on each of these topics is rather limited within the economics of education literature, though other fields of study have provided evidence that each of these factors may impact upon pupil outcomes, thus providing the motivation for each chapter.

It is important to understand how the neighbourhood in which an individual resides may influences their outcomes; the existing literature has identified that neighbourhoods impact upon an array of outcomes, including the probability of school dropout (Overman, 2002; Harding 2003), labour market performance (Manley and Ham, 2010) and teen pregnancy (Harding 2003; Lupton and Kneale 2010). Neighbourhoods may also play a role in determining a pupil's cognitive outcomes; it is evident from UK statistics that pupils from deprived areas perform significantly worse than pupils from non-deprived areas, with a clear and consistent gap in achievement at GCSE level (Department for Education 2014b). While these differentials in educational attainment may be partially explained by individual and family characteristics, the existing literature does suggest that neighbourhoods play a role (Rashbash 2010; Gibbons 2010; Nicoletti and Rabe 2010); however, the evidence is inconclusive, with some studies presenting conflicting results. Difficulties arise in the estimation of neighbourhood effects, thus necessitating the use of advanced econometric techniques; the first chapter of this thesis will adopt a propensity score matching approach, in order to contribute to existing literature and assist in clarifying the currently conflicting evidence. This approach has previously been adopted within the existing neighbourhood effects literature, but has been used infrequently in studies that examine the impact of neighbourhoods upon educational outcomes due to the difficulties that often arise in matching characteristically similar individuals from distinct neighbourhoods. As will be discussed in detail, the data adopted within the chapter facilitates the use of matching methods.

Similarly, school-level policies such as class setting by ability are likely to influence pupil outcomes though pupils and parents may have little control over the pupil's exposure to such a policy. While research outside of the economics of education field suggests that setting has a significant impact upon attainment (Hallam and Parsons 2014; Ireson 1999a), non-cognitive pupil outcomes such as self-esteem and self-concepts are also found to be influenced by setting and ability grouping practises (Abadzi, 1985; Ireson and Hallam, 2009); however, the research that observes non-cognitive outcomes is currently very limited, especially within economics. Few studies have therefore attempted to overcome the methodological issues faced by researchers in the analysis of setting, by using econometric techniques. Concurrently, within the economics of education literature, our understanding of the determinants of non-cognitive outcomes is more limited relative to cognitive outcomes; it is, however, vital to understand the determinants of non-cognitive outcomes due to the fundamental role that such outcomes play in child development and individual life chances (Vignoles and Meschi, 2010). By adopting econometric techniques to analyse the impact of setting on non-cognitive pupil outcomes, the second empirical chapter of this thesis will contribute to the existing literature on the determinants of non-cognitive pupil outcomes alongside the ability grouping research, by analysing whether the school-level policy of setting is beneficial or harmful to pupils' outcomes.

The final chapter of this thesis, which evaluates the impact of converter academies upon pupil outcomes, is motivated by the recent increase in the number of primary converter academies, due to the expansion in the national-level academy programme. Following the election of the Coalition Government in 2010, the academies programme was opened to all primary and secondary schools in England thus allowing all schools to apply to voluntarily convert to a converter academy in order to gain the greater autonomy that academy status offers. At present, there is very little evidence of the impact of converter academies upon pupil outcomes despite thousands of children being affected by the policy; by the end of the 2013/14 academic year, 1,401 primary converter academies existed in England; this number has, and will continue to grow, since following the 2016 budget, the Conservative government expressed their continued commitment to the expansion of the academy programme (Department for Education, 2016). Though the limited existing evidence suggests that academies schools significantly impact upon pupil attainment (Eyles and Machin 2015; Department for Education 2014a), there is no evidence, at present, of the impact upon pupil outcomes at the primary level. The third empirical chapter of this thesis will therefore analyse the impact of primary converter academies upon the cognitive progress of pupils.

In its entirety, the thesis contributes to the economics of education research by addressing the determinants of pupil outcomes outside of the household and classroom, in an attempt to fill the gaps in the existing literature.

Each chapter of this thesis is a micro-econometric investigation based upon pupil-level panel data from a number of different datasets including the Longitudinal Study of Young People in England (LSYPE), the Millennium Cohort Study (MCS) and the National Pupil Database (NPD). A range of methodologies are adopted throughout the thesis, including propensity score matching, fixed effects estimation, instrumental variables and difference-in-difference methodologies, in order to overcome the econometric and evaluation issues met in each chapter. The availability of administrative data sets and the modern econometric techniques, both allow for ability aspects of the pupil to be controlled for, enabling the particular effect of interest in each chapter to be isolated. Pupil outcomes form the common focus of all chapters within this thesis with a range of outcomes being observed, including both cognitive and non-cognitive measures; throughout the thesis, pupils within the compulsory schooling age range and within the UK are the focus of the investigation.

To provide context, the following section will discuss the structure of compulsory schooling in the UK.

1.2 COMPULSORY SCHOOLING IN THE UK

Until September 2013, full-time education was compulsory for all children aged between 5 and 16 in England, Wales and Scotland, with compulsory schooling beginning at the age of 4 in Northern Ireland. From September 2013, upon the completion of secondary school, pupils in England and Wales were required to continue in education until the age of 17; from September 2015, this was raised to the age of 18.

Before beginning compulsory schooling, children under the age of 5 may attend nursery on a non-compulsory, part-time or full-time basis. At the age of 5, children enter the full-time, compulsory primary stage of schooling. The primary education stage accommodates for children between the ages of 5 and 11. Primary education may be split into two stages; infant, known as Key stage 1 (KS1), which caters for children between foundation year and year 2 when pupils are aged between 5 and 7, and junior, referred to as Key stage 2 (KS2), which provides education to children in year 3 to year 6, up to the age of 11. In England specifically, there may be a distinction within some schools between infant and junior stages, since some schools cater solely for the infant or junior stage; other schools simply provide education to all children at the primary stage without providing two separate schools.

At the age of 11, in the final year of the primary education stage, children are required to undertake Standardised Assessment Tests (SATs) that examine pupils in English and maths. Following the removal of the science SATs exam for all pupils in 2009, only pupils within a sample of schools thought to be representative of the whole population are required to take a science test.

Upon completion of primary education at age 11, pupils begin the lower secondary stage of education which provides education to pupils aged between 11 and 16. In a similar manner to the primary stage, secondary education may be divided into two phases; Key stage 3 (KS3), which caters for pupils in year 7 to year 9, when aged between 11 and 14, and Key stage 4 (KS4), which relates to the final two years of lower secondary education when pupils are in year 10 and year 11, and are aged between 15 and 16.

At 16, pupils may stay in full-time education at a school or further education institute; alternatively, until the age of 18, individuals must undertake an apprenticeship or traineeship, or attend part-time education or training while in work.

1.3 STRUCTURE AND CONTENT OF THIS THESIS

Each of the following chapters utilise pupil-level data and adopt econometric techniques to analyse influences upon pupil outcomes in compulsory schooling in the UK.

1.3.1 Brief overview of chapter 2

Chapter 2 analyses the effect of neighbourhood deprivation upon educational attainment at secondary school. Specifically, the chapter attempts to identify whether living in a deprived neighbourhood impacts upon the probability of obtaining five GCSEs A*-C and five GCSEs A* to C including English and maths, known as the gold standard. Data from the Longitudinal Study of Young People in England (LSYPE) is utilised in this chapter which focuses on the impact of neighbourhood deprivation within England specifically.

In order to overcome the issues that surround the analysis of neighbourhood effects, namely selection bias, the evaluation problem and the establishment of causality, the chapter adopts a propensity score matching methodology, allowing for individuals from deprived neighbourhoods to be matched to characteristically similar individuals from non-deprived neighbourhoods.

In addition, the chapter investigates whether there is a differential impact of neighbourhood deprivation according to parental education, to find whether children with parents with post-16 education are influenced to a greater extent by neighbourhood deprivation than children with parents that are not educated to post-16 level. Propensity score matching methods continue to be adopted; the estimated neighbourhood effect for the educated parent and uneducated parent subsamples are then compared, with differences tested for statistical significance.

1.3.2 Brief Overview of chapter 3

Chapter 3 focuses on a school-level policy known as class setting, which involves separating children into classes based on ability in specific subjects. The policy was encouraged by the 1997 Labour government and subsequently widely adopted within primary schools. This chapter looks at whether class setting in maths influences the behaviour of primary aged children.

A fixed effects methodology is initially adopted to identify whether a change in the experience of being set, as opposed to not being set, between the ages of 7 and 11 influences a change in pupil behaviour. Behaviour is measured by the Strengths and Difficulties Questionnaire (SDQ), supplied within the Millennium Cohort Study (MCS) which provides the data for this chapter. The responses of teachers and parents are observed to identify changes in reported behaviour both at home and at school.

This chapter also analyses whether the behaviour of pupils is influenced by the level of the maths set in which the pupil is placed. An instrumental variables approach is implemented in order to overcome the likely endogeneity issue surrounding the level of the set of the child and their behaviour. The methodology involves instrumenting the level of the set with measures of school peer quality.

1.3.3 Brief overview of chapter 4

Chapter 4 examines the impact of primary converter academies upon the progress in pupil outcomes at primary school between KS1 and KS2. This chapter adds to the existing literature by analysing post-2010 primary converter academies, for which, little evidence currently exists.

Adopting 2008-2014 data from the National Pupil Database (NPD), this chapter implements a difference-in-difference methodology to analyse the change in a child's percentile rank within their cohort, according to their average point score (APS) between KS1 and KS2. The rank of treated pupils who experience academy conversion after 2010 and while at primary school is compared with the rank of pupils within the same cohort in the control group. The control group consists of pupils who complete KS2 and therefore primary school, before their school converted to become an academy. This approach defines a suitable control group and therefore assists in overcoming the evaluation problem.

The chapter also attempts to identify whether there is a differential impact of converter academies upon pupils who attend schools that are located within the most deprived neighbourhoods, relative to schools within the least deprived neighbourhoods. The difference-in-difference approach is again implemented with the sample being split according to the deprivation level of the schools' neighbourhood.

CHAPTER 2 : NEIGHBOURHOOD EFFECTS ON EDUCATIONAL ATTAINMENT: DOES FAMILY BACKGROUND INFLUENCE THE RELATIONSHIP?

2.1 INTRODUCTION

To what extent does the quality of the neighbourhood that an individual grows up in influence their outcomes? Empirically, this question has been addressed when considering outcomes such as school dropout (Overman, 2002; Harding 2003), employment prospects and income (Oreopoulos, 2003; Bolster et al. 2007, Manley and Ham, 2010) and teenage pregnancy (Harding, 2003; Lupton and Kneale 2010). One additional outcome of recent interest within the neighbourhood effects literature, and providing the focus of this chapter, is educational attainment.

There is a clear and consistent gap between the educational attainments of young people from deprived neighbourhoods and non-deprived neighbourhoods. The Department for Education (2014b) reported a 29.5 percentage point gap in the attainment of five GCSEs A*-C including English and mathematics in 2012/13, presenting the largest gap at GCSE level between children from deprived and non-deprived areas. Concurrently, it is well documented that children from deprived backgrounds generally surface from school with substantially lower levels of educational attainment (Chowdry, 2010).

Differentials in educational attainment may be explained by a number of factors including individual characteristics, family background and the school attended, alongside neighbourhood characteristics. In an attempt to explain the variability in educational attainment, Rasbash (2010) estimates that a shared environment, including the neighbourhood, but also encompassing primary and secondary schools and the local education authority, accounts for around 22% of the variability of outcomes. The influential factors are, however, unlikely to impact upon educational attainment independently; Cheshire (2007) argues that poor individuals select into poor neighbourhoods, thus factors associated with family background are likely to determine neighbourhood residence. Children from deprived families may therefore also reside within deprived neighbourhoods. Difficulties therefore arise in disentangling the distinct effect of neighbourhood characteristics from other influences upon educational attainment.

This issue in identifying the impact of neighbourhoods upon individual outcomes relates to the problem of a selection bias which may appear in the estimation of neighbourhood effects since it is likely that individuals and families do not randomly select where to live; the choice of neighbourhood may be related to observable and unobservable characteristics of the individual or family. In turn, these characteristics may determine educational attainment, thus a bias arises in the measurement of the impact of neighbourhoods. It is also difficult to recognise causality when distinguishing whether deprived neighbourhoods affect outcomes or whether the characteristics of individuals, which determine neighbourhood residence and are potentially shared by neighbours who also select into the neighbourhood, determine the individual's outcomes. One further problem that researchers of neighbourhood effects face is an evaluation problem; only one outcome may be observed for each individual, hence, neighbourhood effects may not be measured by comparing individual outcomes when living in a deprived neighbourhood to the outcomes of the same individual should they have lived in a non-deprived neighbourhood.

In an attempt to overcome these issues, a number of approaches have been adopted within the neighbourhood effects literature including the observation of sibling and neighbour correlations in outcomes (Lindahl,2011; Nicoletti and Rabe, 2010; Solon et al., 2000), the exploitation of the timing of a neighbourhood move (Weinhardt, 2013) and the observation of a change in neighbourhood composition (Gibbons, 2002; Gibbons et al. 2012) alongside propensity score matching techniques (Harding, 2003), instrumental variable methods (Goux and Maurin, 2007; Cutler and Glaeser, 1997) and analysis of experimental approaches, such as the Moving To Opportunity programme (Sanbonmatsu,2006; Gennetian et al. 2012, Ludwig et al. 2008). Despite the extensive research into neighbourhood effects upon educational attainment, a clear consensus fails to be reached within the literature regarding the magnitude or even the existence, of a role of neighbourhood quality in determining educational attainment. Whereas research within the US provides more clear-cut evidence of neighbourhood effects, studies from Europe and more specifically the UK, reflect much greater variance (Brattbakk and Wessel, 2012).

To correctly identify the effect of neighbourhood deprivation, in an ideal setting the outcome of an individual who experiences neighbourhood deprivation, would be comparable to the outcome of that same individual, should they have lived in a non-deprived neighbourhood. However, due to the evaluation problem this is not possible; only one outcome may be observed for each individual at a point in time. In an attempt to simulate such an experiment, this chapter will adopt propensity score matching methods, allowing for the outcomes of individuals from deprived neighbourhoods to be estimated should they have lived in a non-deprived neighbourhood, by matching characteristically similar individuals. In doing so, this chapter seeks to identify the impact of neighbourhood deprivation upon GCSE outcomes of English pupils, utilising data from the Longitudinal Survey of Young People in England (LSYPE) between 2003 and 2006. The neighbourhood effect estimated will therefore indicate the impact of neighbourhood deprivation, based upon Income Deprivation Affecting Children Index (IDACI) scores, upon GCSE attainment. The GCSE outcomes of interest are the two main GCSE headline measures within the UK: five GCSEs A*-C and five GCSEs A*-C including English and mathematics, also termed the gold standard. With these headline measures gaining much attention as measures of school, programme and intervention success, and being particularly important individual attainments for further and higher education, these measures provide a suitable, observable educational outcome.

Whilst a number of neighbourhood effect studies within the US have attempted to identify the individuals who are more susceptible to neighbourhood effects, for example by investigating the difference in neighbourhood influences between black and white individuals (Harding, 2003; South and Crowder, 1998), there are few studies within the UK that attempt to examine whether neighbourhoods affect all individuals homogenously or whether certain individuals living in deprived neighbourhoods may suffer from the potential negative effects to a greater extent. In addition to identifying the impact of neighbourhood deprivation upon the secondary school educational outcomes on the sample as a whole, this chapter will also consider the role of parental education in determining neighbourhood effects. Specifically, the chapter seeks to identify whether the impact of neighbourhood deprivation differs between young people with educated and uneducated parents therefore allowing for a heterogeneous effect according to family background. Parental education is examined rather than factors such as parental employment status or occupation since such factors may be co-determined with current living arrangements and hence deprivation; parental education, conversely, is likely to be more pre-determined. Parental education is, however, likely to be highly correlated with parental employment status and occupation; the quantity of successful matches of characteristically similar individuals may therefore be limited by the availability of educated individuals living in deprived neighbourhoods. This could be problematic if few individuals are consequently matched between deprived and non-deprived neighbourhoods.

The chapter aims to find whether the differential in outcomes of deprived and non-deprived residents with educated parents is greater than the differential in outcomes for those with uneducated parents; this would be consistent with the hypothesis that the attainments of children with educated parents are improved to a greater extent by living in a non-deprived neighbourhood, relative to individuals with uneducated parents, as this chapter expects to find. Such an effect may potentially be due to the aspirations and ability of the selected peers of those with educated parents differentiating to a greater extent between neighbourhoods, in contrast to children of uneducated parents. Alternative results which would refute the hypothesis of this chapter include findings which indicate that the impact of neighbourhoods impedes more greatly upon those with uneducated parents, thus possibly signalling that family background may compensate or mediate the negative influences. Additionally, an equal neighbourhood effect would insinuate that neighbourhoods have a homogenous influence upon young people with educated and uneducated parents.

This is an interesting question from a policy perspective since the findings may signal a specific group of individuals within deprived neighbourhoods who are more susceptible to the negative influences. Given the findings of this chapter's analysis, it is possible that results may indicate and predict where policies which aim to improve attainment within deprived neighbourhoods will add the greatest value.

The chapter will be structured as follows; a summary of the suggested mechanisms behind neighbourhood effects will be presented in section 2 with a review of the literature which estimates neighbourhood effects following in section 3. A description of the data and the adopted methodology will be discussed in sections 4 and 5 respectively, with section 6 presenting the results from the main analysis alongside the additional models of analysis. This chapter will close with a summary of the chapter aims, methods and results within a conclusion found in section 7.

2.2 BACKGROUND

This chapter attempts to identify whether living in a deprived neighbourhood impacts upon individual outcomes, specifically educational attainments at GCSE level. But why should neighbourhoods matter? There are a number of postulated mechanisms and pathways of neighbourhood effects which explain how neighbourhoods influence an individual's outcomes; though this chapter aims to distinguish whether an overall neighbourhood effect exists, it is useful to identify the possible mechanisms behind this overall effect. Such pathways may then be used to postulate an explanation for the results.

A number of studies have attempted to broadly categorise the mechanisms; Galster (2012 Cited in Van Ham et al. 2012) for example classifies mechanisms into four broad areas including social interactive mechanisms, environmental, geographical and institutional mechanisms. Narrower descriptions of the individual mechanisms and effects will be given whilst adopting this classification.

2.2.1 Social interactive mechanisms

Social interactive mechanisms refer to social processes which are endogenous to the neighbourhood. Collective socialisation describes one of the predominantly deliberated sources of neighbourhood effect which relates to the role models provided by a neighbourhood (Ainsworth, 2002; Galster, 2012 Cited in Van Ham et al. 2012; Ellen and Turner, 1997); young residents may be influenced by the choices, behaviour and lifestyle of other neighbourhood residents. Neighbourhood role models may also convey social norms; should deviant behaviour or in contrast, attainment of higher education be highly prevalent within a neighbourhood, young people may be more likely to adopt or consider such activities. Individuals may see leaving school at age 16, for example, as a norm and attach a lesser stigma, should a high proportion of neighbourhood peers do so (Harding et al., 2010).

Correspondingly, one mechanism through which neighbourhoods may influence outcomes could be through local incentives; with alternatives to formal education being offered by disadvantaged neighbourhoods, for example crime involvement (Lupton, 2006), youths may perceive greater benefits to such activities should local residents or even role models signal a gain. Relatedly, youths may deduce that individuals fail to reap the benefits from responsible behaviour and therefore infer low incentives offered by employment when resident job seekers fail to acquire good opportunities (Ellen and Turner, 1997).

Youths may additionally feel bounded by the achievements of other residents, to whom they relate, thereby influencing their expectations. Galster (2012 Cited in Van Ham et al. 2012) argues such expectations could become self-fulfilling. In contrast, a neighbourhood with a high proportion of higher educated individuals may increase the expectation of entering higher education. The achievements of neighbours may possibly determine an adolescent's perception of available opportunities open to them.

One further postulated social interactive mechanism of neighbourhood effects relates to the monitoring and sanctioning of behaviour. There may be a lesser opportunity for individuals to positively influence youths in neighbourhoods where residents spend less time on positive practises and therefore regulate and sanction deviant behaviour to a lesser extent; this is an issue of social organisation (Ainsworth 2002; Harding et al, 2010). With lesser positive involvement of residents, there may also be a shortage of constructive social networks which possess skills and qualities to assist in education. It is likely that such involvement and opportunities vary between deprived and non-deprived neighbourhoods.

Relatedly, Ellen and Turner (1997) discuss the benefits of a dense social network possessed by an individual since information on employment or even educational opportunities may be more easily obtained with a wide network of friends and family. They argue that individuals with a small social network, or a network that fails to extend beyond the immediate neighbourhood, increases the exposure to the neighbourhood thereby increasing the potential influence of immediate surroundings.

Neighbourhoods are likely to shape an individual's selection of close peers especially when youths reach adolescence and begin to expand their social networks, increasing their exposure to neighbourhood peers. Within empirical work, deviant peer group effects are indicated to influence a youth's grade point average (Darling and Steinberg, 1997 cited in Brooks-Gunn,J., Duncan, G., and Lawrence, J. 1997). Additionally, deviant peers are found to impact upon school attainment, anti-social behaviour and substance abuse (Dubow et al. 1997 cited in Van Ham et al. 2012). Peer ability has also been identified as having a significant impact upon individual educational effort and attainment (Sacerdote, 2001).

2.2.2 Environmental mechanisms

The importance of neighbourhood characteristics is well noted when considering the potential impact upon education, specifically the physical neighbourhood resources such as the provision and quality of local services (Harding, 2010; Lupton, 2006; Ellen and Turner 1997). Facilities such as libraries, local computer facilities and community centres, may impact directly by providing educational resources, whilst provision of other local services, such as healthcare centres, may influence educational outcomes indirectly. Similarly, the distance to post-16 education institutions is found to impact the likelihood of staying on, particularly for individuals on the fringe of participating, such as those from families with low socioeconomic status (Dickerson and McIntosh, 2013). This may additionally influence the perception of academic opportunities open to individuals.

Alongside the social norm and local incentive associations with crime, violence within a neighbourhood may impact upon the education of a young person through other various processes (Harding et al.,2010; Galster (2012 Cited in Van Ham et al. 2012; Ellen and Turner, 1997). Health implications may arise from the stressful situation of living in areas of prevailing violence and crime, for instance anxiety or even injury from being a victim of crime; such health issues may influence school attendance or attentiveness. When adopting US longitudinal data on adolescents' exposure to violence and adult outcomes, Hagan et al. (2001) identify a positive influence on school dropout following exposure to violence of both a verbal and physical nature. Furthermore, results from the Moving to Opportunity experiment indicate a reduced exposure to drugs, gunfire and a decrease in the feeling of unsafeness for individuals who moved from dangerous to lower crime areas. Individuals concurrently reported greater happiness and calmness (Gennetian et al., 2012)

2.2.3 Geographical mechanisms

The geographical location of a neighbourhood may impact upon residents, for example through the proximity to job opportunities; spatial mismatch may surface when opportunities for residents are located far from where they live. An individual's reward perceptions, in regards to education for example, may be manipulated by the availability of local job opportunities (Ellen and Turner, 1997; Lupton, 2006), thus potentially impacting upon educational effort and aspirations.

Furthermore, environmental factors are also suggested to play a role; in a neighbourhood vulnerable to high pollution levels for example, the health of residents may suffer, possibly influencing school attendance.

2.2.4 Institutional mechanisms

Galster (2012 cited in Van Ham et al. 2012) explains institutional mechanisms as the actions typically by individuals outside of the neighbourhood. Stigmatization, may be categorised as an institutional pathway for neighbourhood effects since a stigma attached to a deprived neighbourhood may impede upon residents being offered opportunities from gatekeepers such as employers (Manley at al. 2011); this may apply to employment positions or possibly educational opportunities which may fail to be offered due to the influence of the stigma attached to a neighbourhood and its residents.

2.3 ESTIMATING NEIGHBOURHOOD EFFECTS: LITERATURE REVIEW.

Whilst the mechanisms behind neighbourhood effects and the internal dynamics are important, this chapter concerns itself with the impact of neighbourhoods on individual outcomes and the difference between these outcomes in deprived neighbourhoods relative to non-deprived neighbourhoods. Work of this nature stems from the investigation into the emerging ‘underclass’ within the US in the 1980s; Wilson (1987) identified that poor or disadvantaged neighbourhoods systematically disadvantaged their residents. Development within the neighbourhood effects literature has since seen a growth in strategies and approaches to identify neighbourhood effects and to overcome the main issues surrounding measurement.

2.3.1 Experimental design

Owing to the difficulties in identifying neighbourhood effects, a number of approaches have been adopted within the literature.

As opposed to an observational approach as adopted in this chapter, an experimental design may arguably hold an advantage in measuring the effect of a treatment; an experimental design involves assembling a group of individuals who are equally willing to participate in the experiment or gain treatment. Individuals are randomly assigned to either the treatment group, where individuals are subject to some treatment or intervention, or the control group, where treatment is not received. Random assignment creates two comparable groups of individuals that are statistically equal. The treatment effect may then be identified by comparing the outcomes of individuals who obtain treatment to the outcomes of individuals from the control group who are untreated and are assumed to provide the outcomes of treated individuals should they have not received treatment. The difference in the outcomes between the treatment and control group may be attributed to the treatment effect. Randomised experiments therefore create a treatment and control group when access to a programme is randomly determined; treatment is therefore randomly determined thereby avoiding the selection problem (Bryson et al. 2002). Such a setting thus allows for observation of a neighbourhood effect when isolating the impact of neighbourhood quality upon residents' outcomes.

The Moving to Opportunity (MTO) experiment offers results from such an experimental design. The programme provided the opportunity for families residing in high poverty areas and within subsidized housing, to move to neighbourhoods with lower poverty rates. Between 1994 and 1997, the programme engaged 4,248 families with children under age 18 from developments within five regions of the US[1], randomly assigning participants to one of three groups; a treatment group who received a housing voucher redeemable in the private rental market, a second treatment group who equally received a voucher which could only be used for a rental within a neighbourhood with a poverty rate of lower than ten percent and a third group, the control group, who were not offered housing vouchers and continued to reside in subsidized or public housing. Using the MTO experiment, studies have analysed the impact of neighbourhoods and neighbour quality by comparing the outcomes of treatment participants to those of individuals and families within the control group who possess similar characteristics.

When observing the outcomes of the MTO programme, specifically focusing on educational attainment, studies generally find little evidence of an overall significant influence of neighbourhood quality (Gennetian et al 2012; Kling et al. 2007; Ludwig et al. 2008). Sanbonmatsu et al. (2006) investigate the impact of neighbourhoods on educational achievement when observing reading and math scores of MTO children and comparing these scores between treated and control groups. The study also examines the impact of the programme differentiating between the age of children when participating in the programme, and additionally explores whether the effect varied by gender, race, ethnicity and educational risk. The results fail to convincingly provide evidence of a

[1] Boston, Baltimore, Chicago, Los Angeles, and New York

favourable neighbourhood effect based on the MTO experiment, with insignificant treatment effects identified for combined reading and math scores overall and for all age groups, but also when differentiating by gender and similarly by race and ethnicity.

Chetty et al. (2016) similarly analyse the impact of the MTO programme upon the long-term economic outcomes of children who moved due to the programme. The study aims to identify whether greater exposure to low-poverty neighbourhoods significantly improves outcomes. Relatedly, it is hypothesized that the potential gains from moving to a lower poverty neighbourhood decline as the child's age at the time of the move increases. The study estimates the intention-to-treat (ITT) effects of treatment using an OLS approach; additionally, the treatment on the treated (TOT) impact is estimated by instrumenting treatment take up with treatment assignment since not all individuals offered a voucher to receive treatment took up the offer. The results indicate that children who move to lower-poverty areas before the age of 13 have a higher likelihood of college attendance and higher income attainment while also being more likely to reside in low-poverty neighbourhoods as adults. Contrastingly, for movers over the age of 13, moving to a lower-poverty neighbourhood actually had negative effects upon outcomes; it is suggested that this is due to the effects of disruption. This study therefore presents evidence to suggest that the impact of neighbourhood deprivation may not be homogenous and may in fact vary with exposure.

The neighbourhood effects literature which focuses on experimental data reaches conflicting conclusions. With a similar organisation as the MTO experiment, the Gautreaux programme provided housing vouchers to low-income black individuals who were waiting for public housing in 1981, allowing individuals to move to private apartments; participants were assigned to either a city location within Chicago or a suburban location. Rosenbaum (1995) analyses the impact of the Gautreaux programme on educational outcomes and identifies a significant and positive impact of moving to a better neighbourhood upon high school completion attendance for movers to suburban neighbourhoods, thus providing evidence based upon experimental data that neighbourhood characteristics may influence educational outcomes.

The use of quasi-experimental methods has also been adopted in the estimation of neighbourhood effects (Gould et al. 2011; Aslund 2011), though used more commonly in examining the placement policies of immigrants. Over the period 1987-91, Swedish refugees were randomly allocated to their initial residence by Swedish authorities; Aslund

(2011) exploits the exogenous variation in the original neighbourhood allocation in order to examine the impact upon educational performance of these individuals. The study finds a positive influence of the proportion of highly educated individuals within the neighbourhood upon the grade point average within compulsory schooling.

2.3.2 Non-experimental approach

In the case of neighbourhood effects, there is little opportunity to adopt experimental data for the UK; non-experimental approaches may be alternatively utilized when the issues of concern surrounding the identification of neighbourhood effects are addressed. Non-experimental approaches largely observe whether an individual is influenced by the factor of interest, in this case neighbourhood deprivation, rather than participation as in experimental methods.

A number of alternative strategies have been adopted within the non-experimental literature to overcome the issues surrounding the measurement and observation of neighbourhood effects. Within the literature, the measurement of neighbourhood quality varies, with some studies focusing on the composition of the neighbourhood, such as the characteristics of neighbours, whilst others adopt a similar definition or measurement to this chapter, with a neighbourhood poverty level indicator or deprivation score observed. Cheshire (2007) argues that poor individuals (or their families) self-select, possibly due to affordability, into poor or deprived neighbourhoods; this selection mechanism produces neighbourhoods whose quality or characteristics reflect those of its residents, thus a correlation between household poverty and neighbourhood deprivation is very likely. Estimates of neighbourhood effects relating to quality and neighbour composition or characteristics are therefore likely to be highly associated.

Within the neighbourhood effects literature, social housing is often considered a random source of allocation since individuals do not necessarily have full control over the neighbourhood in which their social housing lies within; residents are therefore to some extent randomly allocated to a neighbourhood. The location of the neighbourhood in which the council house lies within may therefore be uncorrelated with the preferences of the tenant.

Weinhardt (2013) utilizes English data which provides information on social housing residence in order to observe the relationship between neighbourhood deprivation and educational attainment, where deprivation is defined by the social housing density within a neighbourhood. This study exploits the timing of a move into social housing around the time of the national KS3 examinations from 1998 to 2009. The timing of the move is assumed to be exogenously determined due to long waiting lists for social housing within the UK. The study seeks to find whether young people who move home into social housing before the national KS3 exam suffer worse exam scores as a result of the move relative to an individual who moves after the exam. In adopting this strategy of exploiting timing, the study observes a treatment group whose exam results may be impacted by the move before the exam and a control group who move after the exam hence their results are not affected; these individuals are comparable since each are likely to possess unobservable characteristics which are shared by social tenants. When adopting a difference-in-difference framework to analyse the treatment and control outcomes, early movers are identified as underachieving in the KS3 exams, however they did not underachieve to a greater extent than movers after the exams. The findings therefore provide no evidence for short-term negative neighbourhood effects upon educational attainment; movers to social housing tend to achieve similar results regardless of the timing of the move. The results are robust to an alteration of the measurement threshold defining neighbourhood quality, that is, the percentage of social tenants within the neighbourhood.

Gibbons (2002) similarly utilizes social tenancy as a random source of neighbourhood allocation in estimating the impact of neighbourhood composition, measured by the proportion of highly qualified adults, upon the level of education attained by children by the age of 33. The author is able to observe both individuals who are allocated into neighbourhoods by social housing authorities and individuals whose families sort into neighbourhoods based upon the demand for local amenities, together with the corresponding educational outcomes of these individuals within the National Child Development Survey (NCDS). The paper takes two main approaches, initially, exploring the effect of the addition and subtraction of factors within a human capital production function with neighbourhood inputs and then checking for school selection bias by estimating neighbourhood quality. As commonly identified in the neighbourhood effects literature, Gibbons finds that family background matters more for educational attainment than neighbourhood composition. The study does suggest, however, that neighbourhoods impact upon educational outcomes irrespective of family resources with results indicating a greater influence of neighbourhoods on outcomes relative to local school performance factors. In correspondence with Aslund (2011), the study finds that a high proportion of highly educated residents within the neighbourhood is strongly positively correlated with the young person's probability of being highly educated and negatively correlated with their probability of obtaining no qualifications. Ranking neighbourhoods by the proportion of educated adults (with A-levels), adolescents living within a neighbourhood in the top 10% were found to be between 5 and 7 percentage points more likely to attain A-levels than comparable individuals with similar family backgrounds from a neighbourhood ranked in the bottom 10%.

Gibbons et al. (2012) also examine neighbourhood movement but take an alternative approach to Weinhardt's examination of movers by studying the impact of a change in neighbourhood composition upon educational attainment of individuals who stay within the neighbourhood. The study focuses on the impact of residential turnover of similar aged children upon 'stayers' attainment in the KS2 and KS3 UK national exams. Compositional changes examined, include the variation in the proportion of free school meal students and the gender mix, alongside changes in the ability level of neighbours. The paper additionally aims to identify how neighbourhood composition variation impacts test scores across different cohorts. Utilising data from the NPD, the paper estimates within-student differences, obtaining the value-added between the examinations. A changes-in-changes design is therefore adopted in observing the impact of mover-induced neighbourhood composition changes on the educational progression of students through secondary school. Distinguishing the effect of neighbourhoods from school peer effects and controlling for student unobserved characteristics and neighbourhood time-invariant fixed effects, the paper identifies little evidence of significant impact of a change in neighbourhood composition; the effect of neighbourhood peer changes insignificantly influence the KS2-KS3 value added.

Whereas Gibbons (2002) identifies a significant influence of the composition of adult neighbourhood residents, Gibbons et al. (2012) indicate an insignificant role of younger, similar aged neighbourhood peers in determining educational outcomes; this evidence may therefore suggest that adults within the neighbourhood influence to a greater extent than peers, possibly through mechanisms associated with role models or monitoring and sanctioning behaviour for example. However, a study by Gibbons et al. (2014) follows a similar strategy of observing the test scores of neighbourhood stayers to identify potential negative effects of high residential mobility. Findings suggest that neighbourhood turnover of peers does in fact matter for progression of test scores for students between the ages of 11 and 14, though results continue to offer a lack of evidence for an influence of neighbourhood quality upon educational attainment.

Observing sibling correlations is an alternative strategy adopted within the neighbourhood effects literature to overcome the difficulties surrounding measurement; with shared neighbourhood and family background characteristics, siblings present an opportunity to estimate these influences usually by comparing the correlations in outcomes to the correlations between neighbourhood children who share neighbourhood characteristics but differ by family characteristics. Siblings therefore indicate an upper bound of the family background effects whilst the upper boundary of the neighbourhood effect is specified by the neighbourhood peer correlations.

Nicoletti and Rabe (2010) investigate sibling similarities in educational attainment, exploring the relative impacts of the neighbourhood and family background by sibling type. Using NPD data, a decomposition of variance approach is used in comparing educational attainment; the study focuses on attainment at KS2 at the end of primary school and on attainment at GCSE level at the end of secondary school. This approach bounds the effect of both family background and the neighbourhood upon outcomes; subtracting the neighbourhood peer correlation from sibling correlations produces the lower bound of the family effect. The results imply that family background factors have a greater influence on outcomes than the neighbourhood as commonly found within the literature; however, neighbourhoods do play a role in determining educational outcomes, accounting for 9.6% of the variation in pupils' attainment at age 11 and 14.3% of the variation at age 16. Furthermore, neighbourhood effects on attainment in urban areas are greater relative to rural areas, suggestively due to greater peer interaction in urban areas.

Solon et al. (2000) also address the issues surrounding the measurement of neighbourhood effects by observing neighbour correlations. Using the US Panel Survey of Income Dynamics (PSID) dataset to observe individuals aged 25-33 in 1985, the study uses correlations in the later socioeconomic status of unrelated neighbouring children to bound the estimated effect of disparities in neighbourhoods upon the variation in socioeconomic outcomes. Findings of this study correspond with those of Gibbons (2002), Lindahl (2010) and Nicoletti and Rabe (2010), identifying a greater role of family background over and above the influence of neighbourhoods in determining the years of education. The findings do indicate an influence of neighbourhoods, with the proportion of variance in educational outcomes attributable to neighbourhoods equalling 10%, thus indicating a similar effect as identified by Nicoletti and Rabe (2010).

Lindahl (2011) similarly attempts to identify the role of neighbourhoods in explaining the similarities in the income and education of siblings. Using data on individuals born in the Stockholm area in 1953, the study observes correlations in education and income among siblings, neighbouring children and additionally among sixth grade school and class mates, when examining the impact of the school environment. Using maximum likelihood estimation strategies, an upper bound of the neighbourhood effect is estimated by comparing correlations between siblings and neighbouring children. Findings suggest a small influence of neighbourhoods relative to family background factors; the environment explains little of the sibling correlations in education and income outcomes. When observing sibling correlations and class or school mate correlations, it becomes evident that future educational attainment is impacted to a greater extent by the classroom environment than the school or neighbourhood. Nevertheless, overall, neighbourhood correlations are negligible for all outcomes studied for both short-term and long-term consequences with findings consistent with those of Weinhardt (2013).

Propensity score matching has also been used to estimate neighbourhood effects upon a number of outcomes; matching characteristically similar individuals from differing neighbourhoods upon their propensity to be treated assists in attempting to overcome the three main issues in the measurement of neighbourhood effects.

Harding (2003) adopts propensity score matching methods to estimate the effect of neighbourhood poverty on high school dropout and teen pregnancy in the US, using PSID data from 1968 to 1997. Matching children aged 10 on characteristics such as family income, parental education and family structure and estimating the neighbourhood effects upon blacks and non-blacks independently, the study finds that individuals residing in high-poverty neighbourhoods, defined by a poverty rate of over 20%, have a higher probability of high-school dropout and teen pregnancy than comparable individuals in low-poverty neighbourhoods. Estimates of the impact upon high school dropout for example signal that living in a high poverty neighbourhood increases the probability of dropout by around 12 percentage points relative to living in a low poverty neighbourhood; this effect is applicable to blacks and non-blacks alike.

Propensity score matching has also been employed in the investigation of neighbourhood effects upon employment outcomes. Oana and Florent (2012) attempt to identify the influence of living in French priority neighbourhoods upon employment outcomes of school leavers in 2004. The priority neighbourhood label covers three categories of deprived French neighbourhoods. Matching individuals living in a priority neighbourhood to similar individuals living in non-priority neighbourhoods, the study identifies a negative impact of living in priority neighbourhoods upon employment outcomes such as access and quality of employment alongside the acquirement of a full-time employment position. Territorial and residential discrimination are discussed as causes of the detrimental influence of the neighbourhood label upon employment outcomes.

Oana and Florent highlight a potential mechanism of negative neighbourhood effects upon educational attainment; with lower employment outcomes attained by residents, who may also act as role models, the local incentives of employment as projected by residents may be deemed low for young people. These local incentives may directly detract from educational effort, since possibly impacting aspirations. Additionally, the perception of opportunity may be dampened for young residents, potentially prompting low expectations which may become self-fulfilling (Galster, 2012 Cited in Van Ham et al. 2012).

One further approach utilised to estimate neighbourhood effects within the literature is instrumental variables (Goux and Maurin, 2007; Cutler and Glaeser, 1997). Using French data, Goux and Maurin (2007) attempt to identify the impact of close neighbour characteristics on children's outcomes by utilising variation in the proportion of adolescents born at the beginning or end of a year. Since birth dates are likely to be exogenous and an important determinant of school performance, the distribution of neighbour birth dates may be used as an instrumental variable for the neighbours' early educational advancement to examine the impact upon an adolescent's performance at school. The study additionally estimates the reduced-form effect of the family background of neighbours in a poor neighbourhood. This strategy involves the observation of individuals within social housing where assignment is assumed to be quasi-random. The results indicate that neighbours born at the end of the year are more likely to be held back a grade at age 15 than those born at the beginning of the year whilst the performance of children in neighbourhoods with high proportions of neighbours born at the beginning of the year did better than those in areas with high proportions born at the end of the year. The analysis identifies that the probability of being held back is greater when other adolescents within the neighbourhood are also held back. Additionally, when attempting to identify whether a child's school performance is influenced by the level of human capital of neighbouring families, education performance is indicated to be negatively influenced by the share of uneducated families.

Corresponding with the findings of Gibbons et al. (2002) and Aslund (2011), since identifying a relationship between attainment and the educational composition of adults within the neighbourhood, Goux and Maurin additionally recognise an influence of neighbourhood peers upon educational attainment in contrast to the findings of Gibbons et al. (2012). Studies of neighbourhood effects measured by peer ability are comparable to studies of other neighbourhood quality measures; for example, achievement and ability are likely to be largely correlated with socio-economic status (SES), in turn, SES within a neighbourhood is likely to correlate with deprivation and poverty measures.

Owens (2010) looks directly at the SES within the neighbourhood, specifically identifying the impact of relative neighbourhood deprivation, measured by SES, compared with the SES of school peers' neighbourhoods, upon the probability of high school graduation and degree attainment. The study seeks to analyse the impact of neighbourhoods and schools simultaneously using multinomial logit models with interactions between neighbourhood and school characteristics to identify these effects. The results indicate that students from low SES neighbourhoods have a low probability of high school graduation and degree attainment; this probability is reduced further, surprisingly, when the student attends a school with high SES peers. In contrast, the educational attainment of those from high SES neighbourhoods is enhanced by attending a school with high SES peers. Whilst attainment varies by both the neighbourhood and the composition of the school attended, the findings suggest that schools, or higher SES peers, may not compensate for neighbourhood deprivation and may actually exacerbate the effect.

This is an interesting finding to consider within this chapter which will attempt to identify the differential impact of neighbourhood deprivation upon individuals with educated parents relative to uneducated parents. When considering Owens' findings, it may be predicted that individuals with uneducated parents do less well in a higher SES neighbourhood due to the higher SES peers and school composition. Higher SES pupils, or those with educated parents, may then be expected to do much better when living in a non-deprived neighbourhood where they are more likely to attend a school with higher SES pupils. However, whereas Owens considers one aspect of neighbourhood differential, through school peers, the impact estimated within this chapter may signal a multiplicity of neighbourhood mechanisms. Since the results will indicate the difference in outcomes should the same individual have lived in a non-deprived neighbourhood, the neighbourhood effect estimated will ultimately encompass but not entirely reflect only the difference in the school peers of young people between deprived and non-deprived neighbourhoods.

With a similar application to Owens but also this chapter, McCulloch and Joshi (2001) analyse the impact of local concentrations of deprived households upon the cognitive test scores of children aged between 4 and 18. Adopting 1991 NCDS data from the UK, the authors also investigate the role of family characteristics as a mediator of neighbourhood poverty, attempting to identify whether neighbourhood deprivation exerts an impact upon children beyond the influence of families. Neighbourhood deprivation is defined by a number of neighbourhood dimensions identified within the Townsend indicator of deprivation[2]. The results from OLS estimations, where family level characteristics are controlled for, indicate a significant negative influence of family poverty upon test scores for children of all ages. Neighbourhood poverty, however, was found to impede only upon test scores of children aged between 4 and 5 years, independent of other SES indicators. As in many other studies of neighbourhood effects upon educational attainment, the paper concludes that family characteristics are of greater importance than neighbourhoods; in addition, neighbourhood deprivation was not found to be mediated by the physical and social factors associated with the home environment. With regards to the focus of this chapter, the findings of McCulloch and Joshi may suggest that any identified difference in outcomes between children of educated and uneducated parents may not be explained by variations in the home environments. In addition, since this chapter seeks to identify the impact of neighbourhood deprivation upon the educational attainment of young people observed between the ages of 13/14 and 15/16, the findings of McCulloch and Joshi may predict that a neighbourhood effect will not be identified.

Though differing aspects of both neighbourhood characteristics and educational outcomes have been investigated within the empirical research, the neighbourhood characteristics investigated are likely to be largely correlated. Whilst agreement is

---

[2] Proportion of:
1)Labour force unemployed
2)Households with no car access
3)Households with one or more people per room
4)Households not owning their home

reached within the literature in regards to the greater influence of family background factors relative to neighbourhood characteristics, there appears to be a lack of consensus regarding both the magnitude and the existence of an influential role in education of neighbourhoods; this is particularly true for studies within the UK and Europe. This is likely to be largely due to the diversity of methods adopted within the literature to identify neighbourhood effects alongside the variation in the measurement of neighbourhood quality across studies and the outcome of interest. Equally, the level of aggregation at which the neighbourhood is defined, varies across the studies. Within the neighbourhood effects literature that focuses on educational attainment, few papers focus specifically on the same educational outcome, for instance GCSEs. Whilst there is a lack of agreement in the identification of neighbourhood effects, it also seems that there are few methods that are consistently adopted in the measurement of neighbourhood effects.

This chapter will contribute to the existing neighbourhood effects literature by providing an analysis of the impact of neighbourhood deprivation upon educational attainment using the alternative method of propensity score matching to overcome the issues surrounding the measurement of neighbourhood effects. To my knowledge, propensity score matching has not previously been adopted as a method to analyse neighbourhood effects upon educational attainment, though used within studies of neighbourhood effects on school drop-out (Harding, 2003). Whilst adopting an alternative approach to neighbourhood effects measurement, this chapter will concentrate on the impact upon educational attainment at GCSE level specifically concentrating on the attainment of headline measures, that is five GCSEs A*-C and five GCSEs A*-C including English and maths, which assists in filling a gap in the existing UK neighbourhood effects literature where the neighbourhood effect upon such measures remain uncovered. Furthermore, the analysis of the differential impact of neighbourhood deprivation by parental education is an innovative addition to the existing literature, especially for the UK where few studies have attempted to distinguish the characteristics of individuals who are more susceptible to neighbourhood effects.

2.4 DATA

The LSYPE is adopted within this study; this dataset encompasses approximately 15,000 individuals who are followed on an annual basis beginning in 2003/2004 when respondents were aged 13-14 and in year 9 of the UK schooling system. The most recent wave from 2009/2010 corresponds to when respondents were aged 19-20 and had therefore left compulsory schooling. With attrition, by the final wave, the sample size approached 8,700 individuals. Waves one to three will be utilized within this study in order to observe GCSE outcomes corresponding with the year 2005/2006 from wave three when respondents were aged 15-16.

The LSYPE provides a representative sample of young people in England; it is a suitable dataset for the use of this study since it provides a rich source of student information. The LSYPE data were gathered by interviewing the young respondents face-to-face along with one parent or guardian who would be interviewed independently for the initial four waves. Within the surveys, young people were asked to enclose information on personal characteristics, attitudes, personal health, experiences and behaviours alongside information relating to schooling and future plans. In addition, parents were asked to respond to questions concerning the family environment, household characteristics and family background alongside current employment, income and socio-economic status. Information on parental attitudes and educational involvement was also gathered by the parental surveys.

The LSYPE dataset may be matched to the National Pupil Database (NPD) to obtain data on pupils that is unavailable within the LSYPE dataset[3]. The NPD is a longitudinal administrative dataset which tracks all school and college pupils in England throughout their schooling years; it provides detailed information on pupils including prior test scores and exam results alongside pupil characteristics and school information. Matching the LSYPE data to the NPD allows for student past attainments, including KS2 and KS3 test scores, geographical indicators and school level data to be obtained. The LSYPE secure access dataset additionally provided detailed school level information such as the institution type and school historical GCSE attainment. This merging of datasets is important for this chapter in order to match individuals across areas based on ability and school quality.

The LSYPE dataset also provides information on neighbourhood deprivation through the Income Deprivation Affecting Children Index (IDACI),[4] providing a rank alongside a

[3] This data is provided to users pre-matched

[4] The Income Deprivation Affecting Children Index (IDACI) gives the percentage of children under 16 in each lower layer super Output Area (LSOA) who are living with families that are income deprived i.e.

score, which indicates the percentage of children aged under 16 within each lower layer super output area (LSOA) who live within income deprived households; a higher score therefore represents a higher degree of deprivation (Dept. for Communities and Local Government, 2008).The IDACI index is a suitable deprivation measure for this study since it represents the proportion of children directly affected by deprivation within the neighbourhood thereby indicating the deprivation amongst neighbourhood peers and the children observed themselves; however the index is still likely to reflect the characteristics of the adults and the over-16 population within the neighbourhoods given that these individuals determine whether the household is characterised as low income. In addition, since this index is based upon deprivation within the LSOA and around 1,500 individuals are contained in each LSOA, the index provides a suitable measure of deprivation within a small enough area to be defined as a neighbourhood.

The LSYPE secure access dataset provided IDACI deprivation scores for all three waves observed alongside an indicator of home movement thus allowing for an individual's neighbourhood deprivation to be tracked across the observed time period. Within this study, deprivation is defined according to IDACI deciles, with the top 30% deprivation scores characterising the deprivation status of a neighbourhood as being deprived with the rest classified as non-deprived. The definition of neighbourhood deprivation is additionally adopted at the 20% level for comparative and robustness purposes at a later stage within this study.

The LSYPE data set provides information on deprivation scores and individual characteristics from the first wave in 2003 when the respondents are aged around 13 and are in year 9 of the UK schooling system. Many of the characteristics of interest alongside the past neighbourhood residence and corresponding deprivation cannot be observed in previous time periods since this information is not available within the LSYPE. The characteristic controls and deprivation information used within this study therefore correspond with the time period 2003/4-2005/6; the neighbourhood effect presented will consequently indicate the impact of neighbourhood deprivation when exposure duration is at least three years, from year 9 to 11 in the UK schooling system. Kunz et al. (2001) recognise that short-term neighbourhood characteristics are likely to be highly correlated with long-term characteristics thus short-term or point in time neighbourhoods may proxy

---

their families are in receipt of Income Support, Income based Jobseeker's Allowance, Working Families Tax Credit or Disabled Person's Tax Credit below a given threshold.

longer term neighbourhood exposure. The neighbourhood effect estimations may therefore equally signal the impact of longer term neighbourhood deprivation exposure.

Before proceeding with the analysis using LSYPE data, the survey design of the dataset must be accounted for. The sampling process of the LSYPE involved using schools as a primary sampling unit and additionally stratifying on school deprivation levels so that deprived schools were oversampled alongside pupils from ethnic minority groups. Due to this design, the size of the sample from each school is contingent on the school's ethnic composition. This sampling method is applied to attempt to achieve acceptable sample sizes across deprivation levels and ethnic groups in order to have large enough samples to robustly analyse these groups and overcome the issue of non-response. Weighting adjustment is therefore applied. In attempting to avoid under or over-representation of some groups, a skew is introduced within the data. Controlling for the sample weights provided within the LSYPE allows for the panel to be restored, giving representative proportions of respondents from all deprivation levels and ethnic groups (Anders, 2012). The descriptive statistics, presented later within this chapter, will therefore present the raw data once these sampling weights have been controlled for. The sampling weights have not been controlled for within the propensity score matching analysis since it is recommended that sampling weights are ignored with the use of the 'psmatch2' command (Leuven, 2014). This is since sample weights are associated with the characteristics of individuals, which may be directly used in the estimation of the propensity score, or may be highly correlated with these characteristics. The propensity score is then used to match characteristically similar individuals on the basis of their score.

The determinants of attrition between waves one and three were investigated to identify whether those living in deprived neighbourhoods were more likely to 'attrit' from the sample. The results of a logit model with the dependent variable being attrition between waves one and three and the controls being the same characteristics as in the propensity score matching procedure are presented in Table 2.1. Those who leave the sample are significantly more likely to be from deprived neighbourhoods; in addition, 'attriters' are less likely to have parents educated to post-16 level, be born in the UK and be white, have parents with interest in education and attend a mainstream school with a below average A*-C GCSE attainment rate and a class size above average. Attending a school with a below average A*-C rate alone increases the likelihood of attrition. Many of these characteristics are likely to be correlated with living in a deprived neighbourhood, thus it is possible that due to attrition, a smaller sample of individuals within deprived neighbourhoods may be observed in all three periods. The possibility of attrition bias should be considered as, while individuals from deprived neighbourhoods may be more likely to leave the sample, there may also be systematic differences between individuals who possess these characteristics and leave the sample and those with similar characteristics who remain in the sample. Whilst it is useful to identify the characteristics of those who leave the sample to identify any possibility of a bias in results, there is little that can be done to change attrition and problems may occur when attempting to correct for it. Moreover, due to the sampling method of the LSYPE it is likely that individuals possessing the characteristics making attrition more likely were oversampled anyhow.

To test for a neighbourhood effect, the analysis involves observing non-movers who live within deprived or non-deprived neighbourhoods throughout the observation period. Individuals who move between deprived and non-deprived neighbourhoods are consequently dropped from this stage of analysis. It should be noted that individuals may move *within* deprived or non-deprived areas but are dropped when moving *between* neighbourhoods differing by deprivation status. Due to the propensity score matching method adopted, all missing observations must be dropped, thus individuals who have missing values for any of the characteristics of interest in any of the three years are dropped from the analysis.

The initial sample size in wave one equals 15,570, with attrition, waves two and three achieved a sample size of 13,539 and 12,439 individuals respectively. In addition to dropping non-respondents in any of the three observed waves, individuals with missing values of any of the observed covariates are removed from the sample causing a further loss of 2,744 individuals (22% of the initial sample); the greatest proportion of these individuals are lost due to missing data on household employment or parental profession. Further to this, individuals who move between deprivation statuses are removed from analysis, this encompasses 140 observations (1% of the initial sample). The initial sample for analysis therefore includes 9,555 individuals; this represents the sample size used to estimate the full or overall neighbourhood effect. This sample is also used when analysing the neighbourhood effect by parental education, thus missing values and movers are removed prior to the splitting of the sample by parental education. When the sample is defined by individuals with educated parents, the sample size equals 4,621 individuals; for the uneducated parent sample, a sample size of 4,934 is achieved. These figures represent the sample sizes before proceeding with any matching and enforcement of common support which may of course reduce the number of observed individuals.

The primary analysis involves identifying the overall neighbourhood effect when defining a deprived neighbourhood as an area within the top 30% deprived according to IDACI scores. An 'educated' parent is initially defined as at least one parent being educated to at least post-16 level. This definition of an educated parent is adopted since if observing only an educated mother or father, single parent households in which one parent is missing, will be dropped from the sample. It is possible that the effect of mothers' education is distinct from the effect of fathers' education, though the existing literature suggests that both mothers' and fathers' education influence childhood outcomes (Ermisch and Pronzato, 2010); having at least one educated parent allows individuals to benefit from the transmission of abilities, decision making guidance and academic stimulation that educated parents provide. Subsequent analysis will consist of adopting a stricter definition of a deprived neighbourhood with focus on only neighbourhoods with IDACI scores within the top 20%. Additionally, the definition of an educated parent will be varied by defining parents with only a degree or higher as educated as opposed to post-16 education.

2.5 METHODOLOGY

It may be useful at this point to reiterate the study's aim which is to investigate whether neighbourhood deprivation impacts upon educational attainment, specifically focusing on the probability of obtaining 5 GCSEs A*-C and of 5 GCSEs A*-C including English and mathematics, known as the gold standard. Furthermore, the study is concerned with identifying whether the magnitude of the neighbourhood effect is contingent on the educational status of an individual's parents; the impact of neighbourhood deprivation upon individuals with educated parents, with post-16 level education or above, will therefore be estimated and compared with the estimated neighbourhood effect for those with uneducated parents. At this point, the hypothesis of an overall negative neighbourhood effect is posed.

2.5.1 Methodological challenges

As previously mentioned briefly within the introduction, there are a number of methodological challenges in identifying Neighbourhood effects, namely the issues of the evaluation problem and selection bias.

To correctly identify the effect of treatment, that is, a neighbourhood effect, one would like to compare the outcome following treatment of an individual with the outcome of the same individual should they have not received treatment. However, the evaluation problem arises since given the states of either being treated or untreated, an individual may experience one state but not both at any time so only one outcome per person is actually observed, this outcome is known as the factual outcome. Should the individual receive treatment for example, the evaluation problem becomes evident when attempting to observe the counterfactual outcome, that is, the outcome where the individual does not receive treatment. Simply stated, this is a problem of missing data, often termed a problem of causal inference (Holland, 1986), which would be absent should the outcome following treatment and the outcome following non-treatment be directly observable for each individual.

It may be assumed that the treatment effect, in this case the neighbourhood effect, may be observed by comparing the attainment outcomes of those treated with those who are untreated. However, this is likely to give unrealistic estimates; treatment effects may not be homogenous, hence the impact of living in a deprived neighbourhood may differ between individuals.

In the case where treatment effects are homogenous, the average treatment effect (ATE), that is the effect of treatment on a randomly drawn individual, will equal the average treatment effect of the treated (ATT) which is the effect of treatment upon the individuals who are actually treated. With random assignment, the ATE and ATT may be more alike. However, disparity is likely to arise between the ATE and ATT; in a voluntary programme involving treatment of some description, individuals perceived to gain the most from treatment may be the most likely to participate, the ATT would therefore be greater than the ATE and potentially be overestimated if the programme was to be applied more widely to all individuals. This issue remains within this study though a voluntary programme is not observed, this is due to individual’s self-selection (or their parent’s selection) into neighbourhoods and therefore into treatment (Bryson et al. 2002).

Furthermore, the observation of two individuals, such as treated and untreated, rather than the same individual, involves an additional problem; since two individuals are observed, one treated from a deprived neighbourhood and one untreated from a non-deprived neighbourhood, issues arise in the measurement of the treatment impact since the characteristics of the individual may determine their selection into the treatment or control group, whilst to measure the impact, similar characteristics are required between the two individuals to gain an accurate counterfactual outcome for the treated individual. This issue relates to the selection problem.

In order to gain accurate estimates of the treatment effect, self-selection must also be accounted for. Whilst individuals may opt to receive treatment due to perceived greater benefit creating this selection bias, the self-selection problem may also relate to the issue of observable or unobservable characteristics being related to both the treatment and the outcome.

In the case of the calculation of neighbourhood effects, individuals are not likely to randomly choose a neighbourhood, it is more probable that individuals select a neighbourhood in which to reside through residential sorting. One issue is that poor individuals select into poor neighbourhoods (Cheshire, 2007); whilst neighbours are likely to possess similar characteristics to one another, the choice of neighbourhood may be related to an individual's observable or unobservable characteristics. A selection problem may arise when the individual characteristics that are related to the choice of neighbourhood, also influence the young person's educational attainments; this therefore leads to the calculation of a biased neighbourhood effect. Difficulties therefore arise in determining the presence of neighbourhood effects when attempting to control for all factors and characteristics that affect both the educational attainment of a young person and the neighbourhood in which their family lives. Whilst selection bias may be reduced or removed by including all relevant observable variables when the relationship between the treatment and outcome processes may be explained by observable characteristics, the treatment effect will remain biased should the issue not be addressed when unobservable factors influence the relationship, regardless of the number of observable characteristics accounted for within the model.

Relatedly, this selection problem causes difficulties in establishing causality; when observing an individual's outcomes from a deprived neighbourhood, poor outcomes may be attributed to the neighbourhood. However, since individual characteristics are likely to partly determine neighbourhood selection, these characteristics may inevitably lead to poor outcomes despite the characteristics of the neighbourhood of residence. Since residents of deprived neighbourhoods are likely to share similar characteristics, an overall negative neighbourhood effect may be found; thus the identified negative association between neighbourhood deprivation and outcomes may not be causal.

2.5.2 Matching

With non-experimental methods, random assignment does not take place, hence when observing whether individuals were treated or not, self-selection and therefore differences in characteristics between the two groups must be taken into account. The treatment effect may be identified through the procedure and technique of matching as a substitute for randomised experiments (Heckman et al. 1998). Matching methods may take into account the potential self-selection bias on observable characteristics by matching those who receive treatment to individuals in the control group, based upon them having comparable observable characteristics before the treatment is undertaken (Oana and Florent, 2012). Since individuals share characteristics but differ in their neighbourhood deprivation status, the issue of causality may be relieved. Furthermore, matching methods may assist in overcoming the evaluation problem should similar individuals be matched allowing the counterfactual outcome to be observed, this point will be discussed later in more detail.

When referring to the treatment effect, the average treatment effect on the treated (ATT) is specifically the parameter of interest. The ATT indicates the impact of treatment upon those who are actually treated and varies from the average treatment effect (ATE) which indicates the effect of treatment on a randomly selected member of the population.

The effect of neighbourhood deprivation upon educational attainment may be estimated using the ATT:

(Eq. 2.1)

$$ATT = E(Y(1)|D = 1) - E(Y(0)|D = 1)$$

Where Y(1) represents the outcome of interest (GCSE attainment) if the individual is treated and lives in a deprived neighbourhood and Y(0) signifies the equivalent outcome when the individual lives in a non-deprived neighbourhood and is therefore untreated. D indicates whether the individual is treated, and therefore equals one if the young person lived in a deprived neighbourhood throughout the time observed.

The second term is unobserved since it represents the counterfactual outcome for treated individuals should they have lived in a non-deprived neighbourhood, and therefore have been untreated. This term must therefore be substituted by utilizing the outcomes of the control group, identified within the matching approach, to estimate the ATT.

(Eq. 2.2)

$$ATT = E(Y(1)|D = 1) - E(Y(0)|D = 0)$$

The outcome of the control group is therefore used as a counterfactual outcome for treated individuals should they have lived in a non-deprived neighbourhood. This relies upon a number of assumptions, namely surrounding the conditional independence assumption (Caliendo and Kopeinig, 2005).

The matching method relies predominantly on the assumption of conditional independence (CIA); this assumption, also termed *unconfoundedness* (Rosenbaum and Rubin, 1983), states that controlling for observable characteristic differences between the treatment and control groups, where these observable covariates X are unaffected by treatment, possible outcomes, Y, are independent of treatment assignment, that is, the outcome that would result should treatment not be applied would be the same for both groups. To rephrase, conditional on observed covariates X, the outcome Y is independent of the treatment status, denoted by D:

(Eq.2.3)

$$Y(0) \perp D|X$$

The CIA allows for the counterfactual outcome to be assumed equal to the outcome of the control group, hence differences between the treatment and control group outcomes may be inferred as being due to the treatment, thereby indicating the treatment effect. The problem associated with distinguishing causality is therefore alleviated by using this method.

The CIA is violated when all variables influencing simultaneously the treatment and the outcome are not included within the model since the information unavailable to the evaluator will partly account for the treatment effect (Bryson et al.2002).

Ideally, matching should create two groups of identical individuals, when treatment is applied and the untestable assumption of CIA holds, the outcomes of the treatment group will indicate the factual outcome whilst the control group outcomes will present the counterfactual and will therefore represent the outcomes of the treated should they have not received treatment i.e. the outcomes of the individuals from deprived neighbourhoods should they have lived in non-deprived neighbourhoods. Matching methods therefore assist in overcoming the evaluation problem.

2.5.3 Propensity score matching (PSM)

By matching individuals within the treatment and control groups based upon a vector of observable characteristics, the method mimics the feature of randomized experimental data (Heckman et al. 1998). However, unlike experimental data, observational data with matching techniques may not produce treatment and control groups where every individual is matched. In practise, matching individuals on specific characteristics may reduce the quantity of matching possibilities identified as the number of characteristics on which to match increases. This issue was addressed by Rosenbaum and Rubin (1983) who demonstrated the use of a single index, rather than specific characteristics, to match individuals may increase the probability of identifying matches; this index is known as a propensity score.

The propensity score is calculated for each individual within the sample, representing their probability of receiving treatment based upon the observable characteristics specified within the model. With treatment as the dependent variable, covariates within the model should indicate factors which influence both treatment (neighbourhood deprivation) and the outcome variable (GCSE attainment), allowing for a propensity score for each individual to be estimated. The propensity score will therefore indicate an individual's propensity to live within a deprived neighbourhood given these characteristics.

Using the propensity score, individuals from the treatment group may be matched with individuals from the control group on the basis of their propensity score. A number of methods may be used to match individuals, as will be later discussed; the often adopted method of nearest neighbour matching, for example, matches treated individuals with the closest or nearest propensity score in the control group. PSM thereby matches individuals with similar propensities to be treated, and therefore similar propensities to live within a deprived neighbourhood.

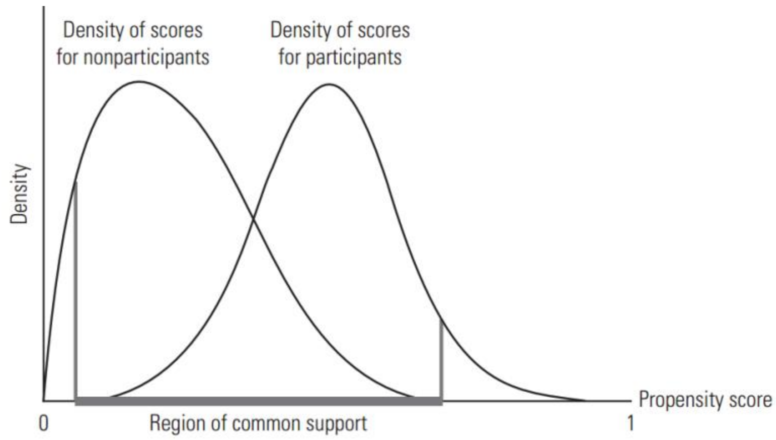
This method of matching therefore assists in minimising the problem of a self-selection bias since individuals who are similarly, though not necessarily equally, likely to live in a deprived neighbourhood are matched whilst this likelihood is determined by their characteristics. Individuals will therefore be similar in terms of their characteristics and their probability to live in a deprived neighbourhood but will differ in terms of the deprivation of the neighbourhood in which they actually reside since a deprived neighbourhood resident will be matched to a characteristically similar individual living in a non-deprived neighbourhood. Since all factors controlling for neighbourhood choice (and educational attainment) are controlled for, this position in the treatment or the alternative control group may be deemed to be random.

Alongside the previously described CIA, the common support assumption additionally underlies matching methods and particularly PSM. The common support assumption states that there must be a positive probability of being both treated and untreated for each individual. This common support or overlap condition therefore ensures that for treated observations there are comparison observations which are close in the propensity score distribution. Observable characteristics that are unaffected by participation, must therefore be similar for the treated and control groups; this may require some untreated observations to be dropped should they be ‘far’ from treated individuals in the propensity score distribution. In some cases, treatment observations may also require dropping should they fail to be matched to a ‘nearby’ untreated observation (Khandker et al. 2010). PSM may therefore present a proportion of individuals for which matches cannot be found.

The common support assumption or overlap condition may be observed visually; Figure 2.1 indicates how the assumption defines a level or region of common support defined by the overlap between the two groups in terms of characteristics. Outside of this region, observable characteristics differ between treated and untreated individuals; hence matching may not be accurate and may therefore be unsuitable. So if common support is attained, there are individuals in the control group close enough to match to a treated individual or there is sufficient overlap in the distribution of treated and untreated individuals.

When undertaking propensity score matching, the common support must be tested, graphical plots to aid visual observation may be found within section 2.5.4

Figure 2.1 Common support region



(*Source:* Khandker et al. 2010).

2.5.4 Propensity score matching application

Applying the propensity score matching procedure to the data involves a number of stages; these steps will now be described in a broader sequence with application to the estimation of neighbourhood effects.

Firstly, a model of location is estimated based upon the pooled sample of individuals. This involves modelling a logit model with treatment, living within a deprived neighbourhood, as the dependent variable. The covariates included within the model should influence or correlate with both the treatment and the outcome so should determine or relate with living in a deprived neighbourhood whilst influencing the GCSE attainment of the young person. Should variables only related with location (neighbourhood deprivation) be included, the outcome, educational attainment, will be unaffected since being unrelated to these factors. Conversely, covariates that are related with the outcome only (educational attainment) shouldn't differ significantly between the treatment and the control group (Bryson et al. 2002). Care should be taken to avoid omission of key factors influencing both treatment and the outcome of interest.

It should be noted that since the logit model estimated is not a determinants model being estimated, the t-statistic and R^2 have little inference (Khandker et al. 2010). Biased estimates will be attained should any determinants of participation be omitted, nonetheless, over-parameterisation should be avoided to avoid higher calculated standard errors for the propensity score.

The logit model will estimate the propensity score of an individual to be treated and therefore their propensity to live within a deprived neighbourhood based upon the specified characteristics.

As previously discussed, the time period observed is 2003/4 to 2005/6 when the young people were aged 13-14 and 15-16 respectively. This corresponds with the individuals being in year 9 to year 11 of the UK schooling system. With data on an individual's neighbourhood residence and characteristics being available from 2003/4, the covariates within the model reflect characteristics which span the three years therefore allowing for the estimation of the individual's propensity score which remains constant over the years observed. For example, a parent is defined as professional should they report holding a professional position for all three waves but non-professional if they do not hold a professional position in at least one of the observed waves; this allows for a propensity score to be calculated which is reflective of the full period observed. When covariates defined in a similar manner to professional parent equal one, they therefore indicate a constant characteristic over the observed time period, when equalling zero, on the other hand, indicates a deficiency of the measured characteristic for a temporary or permanent time, at least for the time period observed. The rationale for this approach is that any changes in characteristics over the time period observed may influence pupil attainment. The neighbourhood effect results will therefore represent the impact of neighbourhood deprivation upon GCSE attainment when exposed to neighbourhood deprivation for at least the three years leading to the examination of GCSEs.

Characteristic controls (X) used to match treated individuals to non-treated include:

- **Household employment**

  This characteristic enters the model as a dummy equalling one if at least one parent is in employment throughout the three waves. Employment may be of either a part-time or a full-time nature. This variable is based on main and second parents' information since this variable is used to match individuals upon the basis of having household income inflows and a working adult within the household. The main or second parents are more suitable for this measure than the mother and father information since the child may not live with either the mother or father.

- **Professional parent**

The professional parent dummy equals one when either of the parents (mother or father) holds 'professional' employment based on the NSSEC (managerial positions) throughout the time observed. The mother and father are used rather than the main and second parent since this variable indicates ability and possibly work ethic; these factors are likely to be initial endowments. As the professional parent may be the mother or father, issues associated with observing a particular parent in a single parent household are overcome.

- **Educated parent**

  Equals one when at least one parent (mother or father) has post-16 education; due to a number of single parent households only one parent is required to be educated. Again, this may signal ability but this may vary from the ability captured by the professional parent variable since professional parents may not necessarily be educated; given the likely age range of LSYPE respondent's parents, it is possible that many worked through the ranks of a company or sector to gain professional employment. This variable also captures potentially having an educated role model and parental interest in education.

- **KS2 ability score**

  A continuous variable giving the average point score from KS2 examinations to match individuals on their ability measured at the age of 10-11.

- **Household deprivation**

  The household deprivation dummy equals one when at least two types of household deprivation are experienced throughout the time observed: no internet access, no computer, no mobile phone, in receipt of free school meals, the household reports financial difficulty.

- **Interaction parental involvement: Parental evening, parental intentions and homework help**

  The interaction equals one when the main parent or partner reports attending parent's evenings in all three periods, the main parent also reports their intentions for the young person to continue in full time education, thus signalling the importance that parents' place upon education, and the young person reports receiving help with homework throughout the time observed. When all three measures hold, the interaction equals one. Whilst indicating involvement, this variable is likely to express the parent's interest and support in the child's education and the importance that the parent places upon their child's education; a combination of each of these three components is likely to express the greatest enthusiasm in the child's education.

- **Interaction: UK born and white**

  This interaction equals one when the young person is both from a white ethnic origin and was born in the UK.

- **School record A*-C**

  The dummy equals one if in 2004 (two years before the young person took their GCSEs) the school that the young person attends did not achieve the 2006 A*-C national average. This measure attempts to control for the quality of the school, with the dummy equalling one for high achieving schools[5].

- **School interaction: School record A*-C, mainstream school and class size above average.**

  This interaction equals one if the school the young person attends has below the 2006 average A*-C rate in 2004, is a mainstream school (i.e. not a special school) and a class size above the average UK rate. This term attempts to also control for school quality and characteristics which may potentially impact on the decision to attend the school and on GCSE outcomes. Controlling for such school characteristics assists in controlling partly for school and neighbourhood overlap whereby bad neighbourhoods may have bad schools, since individuals from deprived and non-deprived neighbourhoods should have similar quality schools, based on these measures if we are to obtain an estimate of the pure effect of the neighbourhood.

A table of summary statistics is provided in Table 2.2 for each of the characteristic controls above including the variable type, mean and standard deviation.

[5] Adopting a dummy variable to control for the achievement record of the school allows for greater balance to be achieved between the treatment and control group in the matching process relative to when school record is treated as a continuous variable. This reasoning also applies to the use of the class size dummy.

Once these characteristics are entered into the logit model and the propensity scores of individuals are estimated, individuals from the treated group may be matched to similar individuals within the control group on the basis of their propensity scores.

A number of matching algorithms may be adopted, each with varying criteria on which to match treated individuals to those in the control group based upon their propensity score. The commonly employed nearest neighbour (NN) matching method will be predominantly adopted within this study with caliper matching additionally employed to check the robustness of NN estimates.

The nearest neighbour matching method involves matching a treated individual with an untreated individual based upon the closest proximity of propensity scores. Hence a treated individual will be matched to an untreated individual who has the most similar propensity to live within a deprived neighbourhood. Multiple neighbours may be used by selecting to match individuals to a number of nearest neighbours. In addition, matching with replacement allows for the same untreated individual to be used as a match for a number of treated individuals since one untreated individual may provide the nearest match for many treated participants. A small number of untreated individuals may therefore provide a close match for many treated individuals; hence, there may be an equal number of matched observations in the treatment and comparison groups yet the treatment group may contain more individuals (Bryson et al. 2002).

The second matching method to be adopted for comparative motives is caliper matching. This is a variant of NN matching which applies a tolerance or threshold for the maximum distance of propensity scores. This method addresses an issue of NN matching whereby the nearest match may not necessarily reflect similar individuals; though being the closest, their propensity may be far from their match. Poor matches may therefore be provided by NN matching. By applying this tolerance, there is a maximum distance that these matches may be apart, by, in essence, applying a caliper which the score must be within, to avoid poor quality matching (Khandker et al. 2010). Setting the width of the caliper usually involves a trade-off between a small sample size, with a small radius, and dissimilar matches, with a larger radius. Within this chapter, a caliper equal to 0.005 is specified for this matching method since the caliper is reduced to the smallest width before the sample size begins to deteriorate.

Following matching, a region of common support must be defined in line with the common support assumption or overlap condition. This region indicates the overlap of the treated and untreated observations and is the only region where the ATT and ATE are defined (Caliendo and Kopeinig, 2005). Observations may need to be dropped should their propensity scores lie outside of this region; common support is defined and enforced within this chapter's analysis by dropping treatment observations whose score is higher than the maximum or lower than the minimum of the score of the controls. Imposing the common support condition leads to no observations being dropped within the matched sample encompassing both educated and uneducated parents; fewer than ten observations fail to satisfy the common support condition and are subsequently dropped within each of the educated parent and uneducated parent samples.

A number of tests of balance may be utilized to check for similarity in characteristics within the area of common support; PSM requires that similar propensity scores are based on similar observable characteristics; when this is true, treatment and comparison groups are said to be balanced. Balancing tests therefore check whether there is equality in the average propensity score and the mean of observable characteristics (Khandker et al. 2010).

Figures 2.2 and 2.3 present the propensity score kernel density plots before and after matching for the main analysis; these figures may be examined to ensure that matching has achieved a balanced distribution of the relevant characteristics collectively across both the control and treatment groups. A visual inspection of these graphs identifies the similarities in propensity score distributions whilst also indicating the area of common support. Given the propensity score on the horizontal axis and the density on the vertical axis, it is clear that before the matching procedure, the two groups, treated and untreated, were dissimilar in terms of the propensity score distributions; accordingly a lack of common overlap is evident. Once individuals in the treatment group are matched to individuals in the control group and individuals without a match are dropped from the analysis, the propensity score densities, as given in Figure 2.3, indicate more correlative propensity score plots with the distribution of scores within the treated group resembling that of the control group. It is apparent that an overlap in the distribution of treated and untreated individuals is achieved by the matching procedure.

In addition to the propensity score plots, the rigorous balancing procedure carried out within the analysis of this chapter also involved implementing a number of alternative balance checks. The results of these may be found in Table 2.3 which gives the results for all three models; the full matched sample, educated parent and uneducated parent samples.

The pseudo R^2 was assessed to evaluate how well the covariates X explain the probability of participation. The R^2 should be low after matching since this signals that no systematic differences exist between the distribution of covariates in the treatment and control groups (Caliendo and Kopeinig, 2005). As can be seen in the first row of Table 2.3, the Pseudo R^2 is very low for each of the specifications.

A t-test balance check allows for the equality of means to be tested between the treatment and control groups. The t-test presents a test of equality between individual covariates between the treatment and control groups after matching. Differences in the means are expected before matching; after matching, it is expected that covariates should be balanced; for adequate balance, differences in the means should be insignificant (Caliendo and Kopeinig, 2005). The final row of Table 2.3 indicates that the educated parent sample and the uneducated parent sample are successfully balanced following the matching procedure. For the full matched sample used to estimate the overall neighbourhood effect, one covariate (the School interaction: School A*-C rate below average, class size above average and mainstream school) has a P value of 0.007, with the treatment group taking a significantly higher value than the control group, indicating significance where we would expect insignificance. There may be slight differences in individuals in the treatment and control group in terms of their school characteristics as represented by the school interaction; as this is an interaction term, one singular characteristic may not explain differences between the two groups since it may be, instead, the combination of these characteristics. However, given that all other balance checks are passed within this sample and this result represents a slight imbalance for a single covariate, this slight discrepancy may be forgiven.

The Hotelling test of equal covariate means similarly tests the distribution of covariates but rather than individually testing the covariates as in the t-test, the Hotelling test checks for the joint significance of covariates. The test, which follows the t distribution, should indicate insignificance should balance be attained. From the second row of Table 2.3, it is evident that according to the Hotelling test, balance has been achieved in all three samples.

Additionally, the standardised bias check is carried out. Due to non-randomized assignment, a bias is likely to arise due to the self-selection, as previously discussed; this check tests the extent to which the bias has been reduced by PSM. The standardised bias gives the percentage difference in the sample means in the treated and control group samples as a percentage of the square root of the average of the sample variances in both groups. There is consensus that a standardised bias reduction to below 5% after matching is considered sufficient, thus on the basis of the results, the standardised balance test suggests that balance is achieved since the bias is highest within the educated parent sub-sample but equals 3.3%.

Similarly, the absolute bias should be reduced for individual covariates by the matching procedure; the absolute bias indicates the absolute value of difference in the mean value between the treatment group and control group. Though a bias below 5% is preferable, Grilli and Rampichini (2011) express that a small unbalance, such as 11% as in their example, is acceptable. The absolute bias, though marginally greater than 5%, does indicate successful balance for each sample.

In addition, it should be considered that the model and specification adopted was required to satisfy the balancing checks for all three samples.

Whilst the balance is tested and deemed important, emphasis is placed upon the use of a common specification across all three samples; the consistency of the controls and the model provide a good basis for analysis and comparability across all samples where the specification managed to achieve balance in each individual sample. This rationale to gain a common specification whilst attaining good balance across all three samples explains why some variables are not included; the common specification of all three samples achieves a good level of balance whilst controlling for a number of important determinants of both educational attainment and neighbourhood deprivation.

Figure 2.2 Propensity score plot: before matching

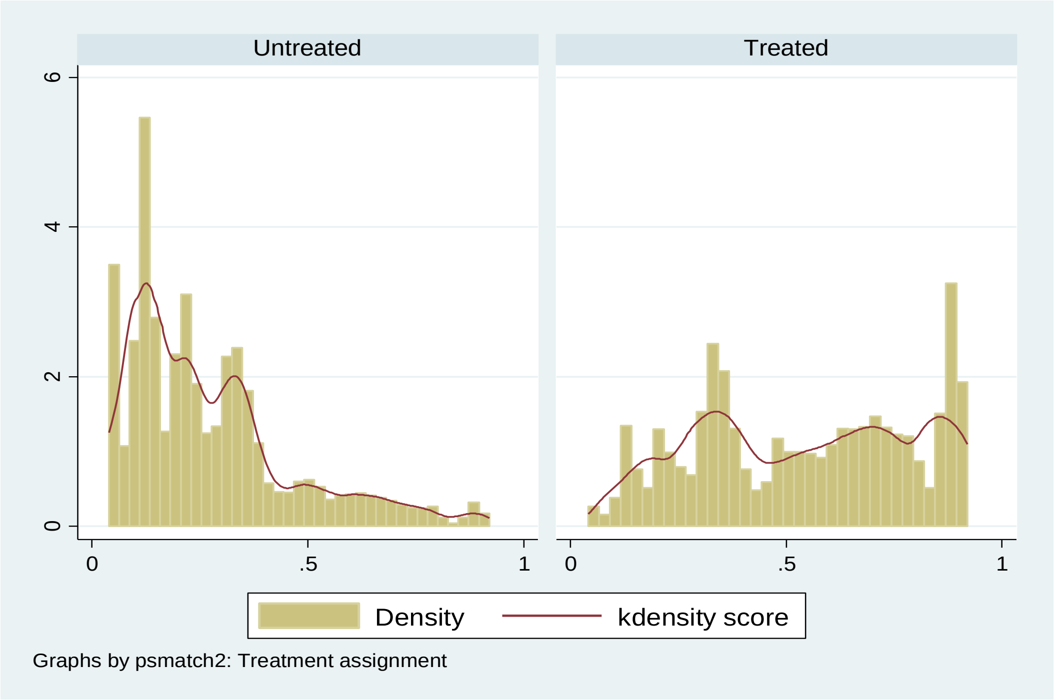
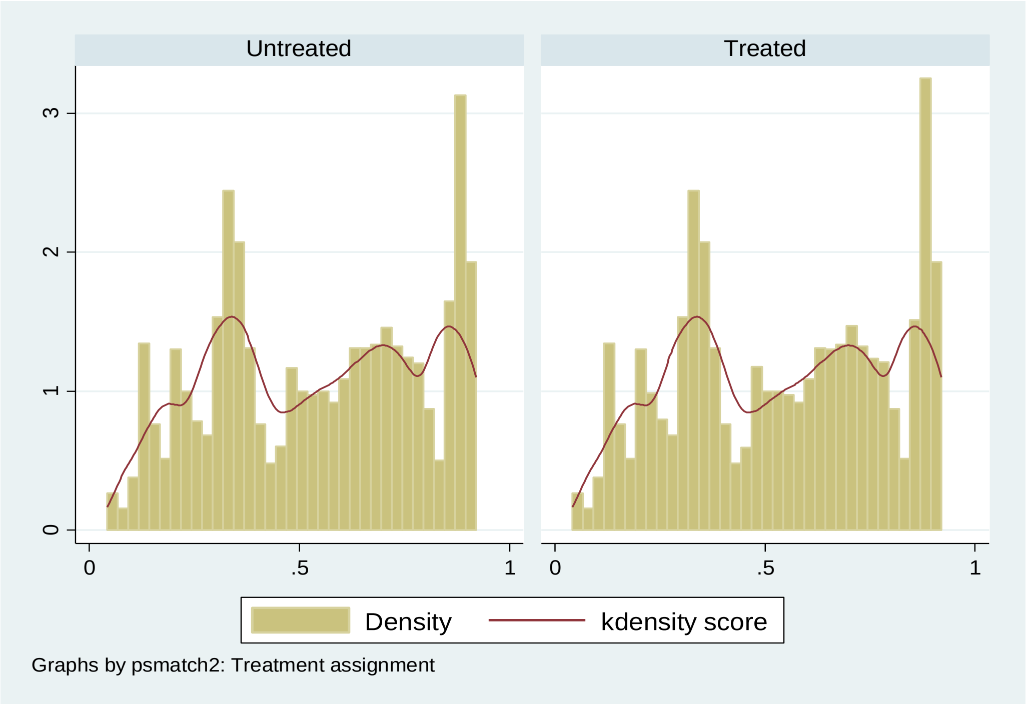


Figure 2.3 Propensity score plot after matching



The standard errors obtained and presented within this analysis were acquired by bootstrapping. Prior to the matching procedure, propensity scores are estimated, they are not known; this estimation is likely to involve some variance which should be included within the estimated variance of the treatment effect. Due to the estimation stages, variation is likely to exceed the normal sampling variation whilst the standard errors are likely to be undervalued. Bootstrapping provides a resolution to this issue; the bootstrapping procedure involves repeatedly estimating properties, such as standard errors and bias, from samples that are drawn from the original full sample. Alongside the results, the first steps of estimation, including propensity score and common support, are re-estimated for each bootstrap draw. The number of times the bootstrapping procedure is repeated equals the number of bootstrap samples and the number of estimated average treatment effects. The sampling distribution may then be approximated by the distribution of these means (Caliendo and Kopeinig, 2005).

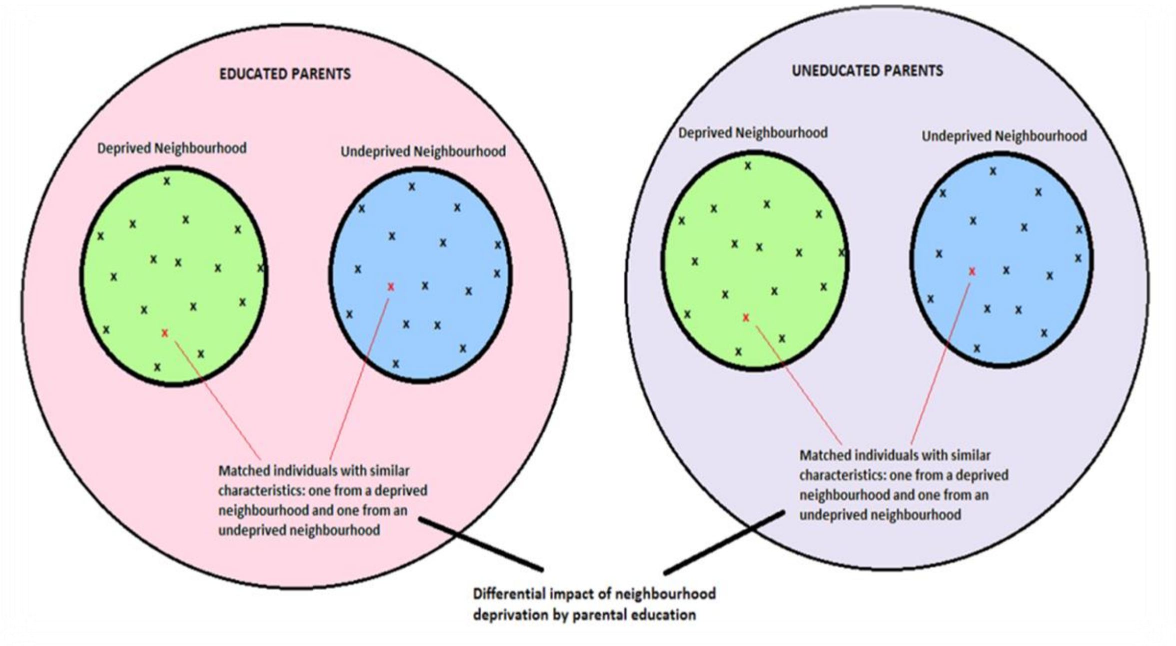
2.5.5 Methodology part two: Observing neighbourhood effects by parental education level

The propensity score procedure discussed above explains the methodology adopted for the first part of this paper. For the second part, the study aims to address whether a subset of the sample are more impacted by treatment than the overall sample; that is, are children from uneducated families more susceptible to neighbourhood effects than the children of educated families?

In order to address this question, propensity score matching techniques will continue to be adopted with the procedure explained following an identical arrangement. However, before estimating an individual’s propensity score and matching based on this score, the sample is split according to parental education. Propensity score analysis will be carried out on the two separate groups to identify the neighbourhood effect for individuals with an educated family background (or at least one educated parent) alongside the neighbourhood effect for those individuals with uneducated parent/s. In doing so, individuals with educated parents are matched to others with educated backgrounds, differing on their neighbourhood deprivation; education will therefore be over-weighted so that individuals are matched exactly on this characteristic whilst other covariates are treated unequally relative to family education; the remaining previously matched characteristics continue to be accounted for within the propensity score. A visual aid to this explanation is provided in Figure 2.4.

The neighbourhood effect will be calculated as before with GCSE outcomes of those in a deprived neighbourhood compared with the outcomes of those living in a non-deprived neighbourhood, yet this will be calculated twice; once for those with educated parents and independently, for those with uneducated parents. The average treatment effect on the treated (ATT) will be compared between sub groups. From this strategy, a higher treatment, or neighbourhood effect identified from the individuals with educated families may be concluded to indicate a greater differential influence of neighbourhoods upon those from educated backgrounds. An equal effect and therefore zero difference between educated and uneducated parents' children's outcomes would imply that family background, in terms of education, does not alter the impact of neighbourhood deprivation upon young people's outcomes.

Figure 2.4 Observing neighbourhood effects by parental education



2.6 RESULTS

2.6.1 Descriptive Statistics

To investigate the impact of neighbourhood deprivation upon educational attainment, the outcomes observed include whether the individual gains five GCSEs graded A* to C, with the additional outcome of whether the individual gains the gold standard of GCSE results, that is five GCSEs A*-C including English and mathematics. The treatment for this initial part of analysis refers to living within a deprived neighbourhood, defined in the top 30% deprived by the IDACI score, for all three years observed between 2003-2006. This section discusses the raw data before performing propensity score matching and providing a formal analysis of results.

Table 2.4 gives the raw percentages of individuals attaining the GCSE outcomes of interest within deprived and non-deprived neighbourhoods. It is clear that the attainment rate of both outcomes is higher in non-deprived neighbourhoods relative to deprived, for example, 41.9% of residents in deprived neighbourhoods obtain 5 GCSEs A*-C relative to 66.7% in non-deprived neighbourhoods. The achievement of the gold standard is lower within both deprived and non-deprived neighbourhoods at 28.7% and 55.7% respectively.

This is also evident when observing attainment by neighbourhood and by parental education. As given in Table 2.5, the raw effects indicate that the percentages gaining the GCSE outcomes are higher in non-deprived areas for both attainment outcomes and for both levels of parental education; for example, within deprived neighbourhoods 56.1% of those of educated parents, defined as those with post-16 education, attain five A*-C relative to 76.7% of children of educated parents in non-deprived neighbourhoods. Similarly 40.3% of individuals in non-deprived neighbourhoods with uneducated parents obtain 5 GCSEs A*-C including English and maths, compared to just 22.8% of deprived neighbourhood residents who have uneducated parents. It is evident that the raw gaps in attainment between children in deprived and non-deprived neighbourhoods are greater amongst children of educated parents both in the attainment of five GCSEs A*-C and the gold standard; the attainment gaps between deprived and non-deprived residents with uneducated parents are markedly smaller. These raw attainment gaps are greatest when observing the attainment of five GCSEs A*-C including English and maths; this gap between deprived and non-deprived neighbourhood equals 23.8 percentage points for individuals with educated parents, compared to a 17.5 percentage points gap for individuals with uneducated parents.

In addition, the statistics given in Table 2.5 indicate a higher attainment among children with educated parents relative to uneducated parents within both deprived and non-deprived neighbourhoods.

Table 2.6 presents the raw data associated with the secondary part of analysis when the neighbourhood effect will be analysed according to the level of parental education. Within the overall sample, 49.9% of young people have parents educated to post-16 level thereby representing close to half of the sample. A much smaller proportion of parents possess a degree with 15.4% educated to this level.

The raw data on parental education may be linked to data on a young person's neighbourhood deprivation to provide an indication of the proportions of educated and uneducated parents within deprived and non-deprived neighbourhoods but also to locate the neighbourhood residence, by deprivation, of the educated parents within the sample.

As indicated in Table 2.7, within the deprived neighbourhoods, only 30.2% of the young people observed have parents who are educated to post-16 level, whilst 69.8% have uneducated parents. The ratio of educated to uneducated parents is much higher within non-deprived neighbourhoods with 59% of young people within the sample having parents who are educated.

Table 2.8 shows the proportion of the educated sample who reside within deprived and non-deprived neighbourhoods respectively. Of the individuals with educated parents, only 19.1% reside in deprived neighbourhoods. These small proportions of educated parents residing in deprived areas indicate somewhat the selection issue whereby 'poor' or individuals from low-socio economic backgrounds select into 'poor' neighbourhoods since over 80% of parents with a post-16 education may be located within a non-deprived neighbourhood. These figures additionally highlight the potential problem of gaining sufficient data on educated individuals within a deprived neighbourhood in order to perform matching methods.

2.6.2 Propensity score matching estimates

Table 2.9 presents the main results of the evaluation. The neighbourhood effect for the full sample, given in the first row, presents the overall effect of residing within a deprived neighbourhood, defined by being within the top 30% of IDACI scores, for at least the three year time period observed. This effect, which is estimated by the propensity score matching procedure, is given by the difference in outcomes of the treatment and control group and indicates the impact upon GCSE outcomes of living within a deprived neighbourhood. Nearest neighbour and caliper matching estimates are given with the estimated effects from the procedures indicating that results are robust to a change in the matching procedure. As very similar estimates are obtained by the two procedures, the nearest neighbour effects will be discussed.

The first panel looks at the impact of neighbourhood deprivation upon the attainment of five GCSEs graded A* to C. Ceteris paribus, the results indicate that individuals within a deprived neighbourhood are 4 percentage points less likely to achieve these GCSE grades than comparable individuals within the control group who live in a non-deprived neighbourhood; this is a significant effect at the 10% level of significance. Given that 66.7% of non-deprived neighbourhood residents achieve five GCSEs A*-C, comparable with 41.9% of deprived neighbourhood residents, the estimated neighbourhood effect may explain 16.1% of the raw gap in attainment of five GCSEs A*-C between deprived and non-deprived neighbourhood residents. This finding indicates that neighbourhoods do matter in determining educational outcomes of young people thus conflicting much of the neighbourhood effects literature which finds a zero or very small effect of neighbourhoods upon individual outcomes. This variation in results between this study and other papers alongside within the neighbourhood effects literature may be explained by the disparity in a number of factors such as the methodology implemented, the outcome variable or measure of interest, the data adopted alongside possibly the definitions of a neighbourhood or deprivation.

Continuing with the analysis of the overall effect of neighbourhood deprivation, the focus now turns to the outcome of attaining five GCSEs A*-C including English and maths, termed the gold standard. As indicated in column 5, young people living in deprived neighbourhoods are 6 percentage points less likely to attain the gold standard of GCSE results relative to a similar young person who lives in a non-deprived neighbourhood, ceteris paribus. This effect is significant and suggests that neighbourhoods do partly determine the GCSE outcomes of young people when we additionally consider whether good grades in both English and mathematics were attained. Considering that 28.7% of individuals living in deprived neighbourhoods within the sample attain at least five A*-C grades including English and mathematics, relative to the 55.7% in non-deprived neighbourhoods, this estimated effect suggests a sizeable impact of neighbourhood deprivation upon young people's outcomes; neighbourhood deprivation explains approximately 22.2% of the gap in the attainment of the gold standard between deprived and non-deprived neighbourhood residents. Again, this finding contrasts with several neighbourhood effects studies.

The findings suggest that neighbourhoods play a greater role in determining whether an individual attains 5 GCSEs A*-C including English and mathematics, than in influencing the achievement of any 5 GCSEs with good grades. The reason for this may be that individuals whose educational attainments may be suffering from the mechanisms and effects of neighbourhood deprivation could possibly fail all GCSEs except a small number so the individual may enter the five A*-C outcome category. However, attaining good grades in at least five subjects including the core subjects, English and mathematics, signals an array of skills and abilities in subjects. Students are likely to understand the importance of good GCSE attainment within English and mathematics so presumably aim to achieve good grades in these subjects; it may therefore not be a matter of effort but underlying characteristics and factors, such as neighbourhood effects which influence this outcome. For these reasons, the results are as expected: neighbourhood deprivation has a larger influence on the attainment of an arguably more difficult set of GCSE results with greater importance for future prospects.

Though the significant negative neighbourhood effect identified contrasts with a proportion of the neighbourhood effects literature, the findings do correspond with those of Nicoletti and Rabe (2010) who identify that 14.3% of the variation in pupil attainment at age 16 in England may be attributed to neighbourhoods, when considering GCSE outcomes in the core subjects. The greater neighbourhood effect identified, relative to Nicoletti and Rabe, may be explained by the differential focus in GCSE outcomes; whereas Nicoletti and Rabe estimate the impact upon the GCSE score, which is associated with the grades received in the observed core subjects, this chapter looks at a binary outcome thus observing whether an individual attains any five GCSEs graded A* to C or not. One higher core subject result is able to compensate for a low outcome in a different core subject within the Nicoletti and Rabe measure, whereas the GCSE grades of core subjects are considered individually within the five GCSE A*-C including English and maths measure adopted within this chapter.

Table 2.9 also presents the estimated effect of neighbourhood deprivation by parental background, giving the impact of neighbourhoods upon individuals with parents educated to at least post-16 level, alongside the effect upon individuals with uneducated parents with their highest level of education being below post-16 level. This analysis seeks to identify whether individuals with educated parents incur a differential neighbourhood effect relative to those with uneducated parents. A neighbourhood effect equal to zero for any estimate would imply that when living in a deprived neighbourhood, the likelihood of obtaining the GCSE outcomes is not different to the likelihood of those living in non-deprived neighbourhoods achieving these outcomes. When observing the distinct neighbourhood effects for educated parents and uneducated parents, a difference of zero, or a difference that is insignificantly different from zero, between the effect upon educated and uneducated groups would imply that parental education does not alter the influence of neighbourhood deprivation upon the child's attainment at GCSE level.

From column 1, the neighbourhood effect upon the GCSE attainment of those with educated parents is found to be negative; ceteris paribus, individuals with educated parents living within a deprived neighbourhood are 7.4 percentage points less likely to attain five GCSEs graded A*-C than similar individuals with educated parents living within a non-deprived neighbourhood. This effect is significant at the five percent significance level and indicates that children with educated parents do less well in terms of the GCSE attainment of five good grades, than children who possess comparable characteristics but live in a non-deprived neighbourhood. This is a sizeable effect if we consider the raw data; 76.7% of individuals living in a non-deprived neighbourhood with parents educated to at least post-16 level attain five GCSEs A*-C; this is comparable with 56.1% who attain these grades in deprived neighbourhoods. The true neighbourhood effect therefore seems to equal 35.9% of the raw attainment differential between deprived and non-deprived neighbourhoods.

Correspondingly, this effect is calculated for individuals with uneducated parents; ceteris paribus, estimates reveal that young people with uneducated parents living within deprived neighbourhoods are 1.7 percentage points less likely to attain five GCSEs graded A*-C than similar individuals who live within a non-deprived neighbourhood. However, neighbourhood deprivation does not significantly influence the attainment of five A*-C for individuals with uneducated parents.

Comparing these results highlights a greater influence of neighbourhoods upon the educational outcomes of those with educated parents relative to individuals with uneducated parents. From the difference given in column 2, there is a 5.7 percentage point difference between the estimated neighbourhood effects for the two groups. This difference is insignificant suggesting that there is not a significant difference in the impact of neighbourhood deprivation upon the attainment of five GCSEs A*-C between individuals with educated and uneducated parents.

Turning attention to the gold standard of GCSE attainments, the neighbourhood effect is again calculated and compared between individuals with educated and uneducated parents. Ceteris paribus, individuals from an educated background living within a deprived neighbourhood are 12.3 percentage points less likely to attain at least five GCSEs A*-C including English and maths relative to similar individuals in the sample with educated parents who live in non-deprived neighbourhoods. This is a highly significant impact of neighbourhood deprivation, indicating that children of educated parents could do much better should they have lived in a non-deprived area. When adopting caliper matching methods, this estimated effect is slightly higher equal to 12.8 percentage points. Given the NN estimates, the results indicate that neighbourhood deprivation explains 51.7% of the raw gap in the attainment of the gold standard GCSE results of children with educated parents from deprived and non-deprived neighbourhoods.

Similarly, the estimate of the neighbourhood effect upon children of uneducated parents indicates that those living in deprived neighbourhoods are 5.7 percentage points less likely to attain the gold standard GCSE result relative to people from non-deprived neighbourhoods, ceteris paribus. This effect is also significant at the five percent significance level, explaining 32.6% of the raw gap in the gold standard attainment between children with uneducated parents living in deprived and non-deprived neighbourhoods.

Individually, each of these effects is greater than the impact identified when observing the five A*-C outcome, suggesting that neighbourhoods influence the probability of attainment of good GCSE grades including English and maths to a greater extent than the probability of gaining any five GCSE graded A*-C, as expected. As argued previously, this may be due to the general attainment of five A*-Cs being less demanding than gaining good grades in English and maths also; a negative impact of deprivation may therefore still allow for some GCSEs to be attained whilst impeding upon the achievement within the more difficult or vital subjects.

If the educated and uneducated neighbourhood effects are compared and the difference is calculated, it is clear that neighbourhoods influence the outcomes of the educated group to a greater extent than the uneducated group. The impact of neighbourhood deprivation is 6.7 percentage points greater for those with educated parents relative to those with uneducated parents. This significant finding suggests that the losses, in terms of educational outcomes, from living in a deprived neighbourhood are greater for those with educated parents relative to uneducated parents. To rephrase, the difference between what individuals with educated parents attained in deprived neighbourhoods and what they could have attained should they have lived in a non-deprived neighbourhood is significantly greater than the difference between actual achievement in deprived areas and potential attainment in non-deprived areas for individuals with uneducated parents. This is almost a value-added effect where the attainment of the uneducated parent group differs slightly between neighbourhood deprivation status whereas children of the educated add much more value, in terms of the probability of attaining five GCSEs A*-C including English and maths, when living in a non-deprived neighbourhood rather than a deprived neighbourhood.

From these results, it is not true that children from educated parents do worse than those from uneducated parents, in fact, the underlying ability of the educated is likely to be higher than that of the uneducated group given their parents' ability. Raw statistics from Table 2.5 indicate greater proportions of individuals with educated parents attaining the two GCSE outcomes relative to those with uneducated parents; this is true within both deprived and non-deprived neighbourhoods. What the results do suggest is that the educated group in deprived neighbourhoods could have had a better chance at attaining the gold standard if they had lived in a non-deprived neighbourhood. The potential gain in the likelihood of gaining the gold standard for the uneducated if they had lived in a non-deprived neighbourhood is significantly lower.

The explanations behind these results are based on speculation alone; the results may correspond somewhat with Owens (2010) who identifies low SES neighbourhood children as being worse off when attending schools with a high composition of high SES children, whilst high SES pupils do better by attending such schools. Owens observes the impact of attending a school which may be more typical of a non-deprived area whilst this study considers a difference in neighbourhoods thus encompassing a number of mechanisms and effects possibly including such school effects. Applying these findings, it may therefore be that, as Owens identifies, low SES individuals or even the children of the uneducated parents do worse in high SES schools, possibly within non-deprived neighbourhoods, however, other mechanisms and effects of the non-deprived neighbourhood positively influence attainment as expected, since living in a non-deprived neighbourhood does improve outcomes overall. High SES children or those of educated parents may then be positively influenced by both the school alongside other neighbourhood characteristics when moving from a deprived to a non-deprived neighbourhood. The effect of moving to a school with a higher SES composition and moving to a less deprived neighbourhood may therefore work in the opposite direction for those with uneducated parents but in the same direction for children of educated parents.

An alternative view on this result may be that this neighbourhood effect, which widens the gap between deprived and non-deprived residents, does so to a greater extent for individuals from educated parents than uneducated. This may be due to children of educated parents being more greatly affected by having peers, such as friendship groups, classmates or school peers who are of a lower ability or lower socio-economic background, as within deprived neighbourhoods, than children of uneducated parents who may potentially be of a lower average ability. Research does suggest that higher ability students are more sensitive to school composition (Opdenakker and Van Damme, 2001). It could be, therefore, that children of educated parents who have a higher level of innate ability are more sensitive to being within a school or classroom with a high proportion of low ability students which may be a feature of schools within deprived neighbourhoods. Children of educated parents may therefore be more negatively influenced by deprived neighbourhoods than those of uneducated parents.

Evidence advocates that low ability peers may impede upon the results of students including those of higher ability students within a school environment (Lavy et al.,2011). Lavy et al. (2012) define ‘bad’ peers as those within the bottom 5% of the ability distribution and identify that reducing the proportion of bad peers from 20% to 0% increases the age 14 test scores by 0.17 of the within-pupil standard deviation in the distribution of these test scores. Moreover, this study fails to identify evidence of good peers, in the top of the ability distribution, influencing attainment. It could be argued that the children of uneducated parents living in deprived neighbourhoods are more likely to be ‘bad peers’ themselves relative to the children of educated parents; a higher proportion of children with educated parents within deprived neighbourhoods may therefore be affected by the low ability bad peers within deprived neighbourhoods whereas, since the children of uneducated parents are likely to be of lower ability, they are therefore more probable to be the bad peers themselves, thus a lower impact of this ‘bad peer’ neighbourhood characteristic for the uneducated group is observed. This may then explain a larger overall negative influence of neighbourhoods upon the group of children with educated parents.

Another possible explanation of this identified effect may be due to the impact of peer aspirations and attitudes rather than, or in addition to, peer ability. A young person's aspirations to attend post 16 or higher education may be correlated with the aspirations of their friends or close peers; Alexander and Campbell (1964) identify that male senior pupils are more likely to attend, aspire to attend and expect to attend college when their best friend does relative to when having a best friend who does not intend to attend college. Furthermore, aspirations are found to impact upon educational outcomes (Ryan and Homel, 2014). Since lower socio-economic backgrounds and low income influence lower aspirations of young people relative to more advantaged peers (Schoon, 2006), it is likely that the average aspirations to continue in education or to do well in education are lower amongst peers in deprived neighbourhoods where a higher proportion of low SES families reside. Christofides et al (2012) identify that the effect of peers goes beyond the influence of parents, teachers and school characteristics. Moving from a deprived neighbourhood, where educational aspirations to stay on or achieve good results for example may be low, to non-deprived neighbourhoods where aspirations among peers may be higher, may therefore increase aspirations and achievement levels for both individuals with educated and uneducated parents. However, due to the relationship between SES and aspirations, this impact of improved peer aspirations may be bounded for those with uneducated parents so that aspirations do not increase to such a great extent as individuals with educated parents by having peers who have higher ambitions. This may be due to a missing role model within the family or a lack of information for those with uneducated parents.

One further possible explanation, again based purely on conjecture, may be that a move from a deprived neighbourhood to a non-deprived neighbourhood could encompass other lifestyle variations which may differ between educated and uneducated families. For example, Lupton (2003) argues that the social relations of individuals will vary between isolated and well-connected areas; within non-deprived neighbourhoods, educated parents and their children alike may have a greater opportunity to expand and build social networks with other educated individuals and families therefore possibly increasing the exposure to potential educated role models. Young people with educated parents may associate with peers with more similar characteristics; with educated families being underrepresented within deprived neighbourhoods, the young person's opportunity to associate with individuals of similar backgrounds may be limited. Uneducated parents and their children may, on the other hand, continue to associate with individuals of similar backgrounds and socio-economic status, as they may have done when living in a deprived neighbourhood, hence the social networks of young people with uneducated parents within deprived neighbourhoods may be very similar in non-deprived neighbourhoods, whereas the networks of individuals with educated parents may vary between deprived and non-deprived neighbourhoods. With social networks and peers influencing upon attainment, this difference in friends and associates between deprived and non-deprived neighbourhoods may explain the greater neighbourhood effect for the educated parent group.

Furthermore, characteristically similar individuals in non-deprived areas may lead differential lifestyles to those in deprived neighbourhoods, thus impacting upon educational attainment. For example, extracurricular activities are found to enhance educational and occupational aspirations (Gutman and Akerman, 2008). Xu et al. (2009) identify a negative influence of neighbourhood disadvantage upon the participation in extra-curricular activities whilst those with educated parents are more likely to participate. There may therefore be little difference in the participation in such activities between deprived and non-deprived residents with uneducated parents, whereas the participation of those with educated parents in non-deprived areas may be greater than the participation of individuals with educated parents in deprived neighbourhoods. Individuals with educated parents may therefore benefit from involvement in such activities through other mechanisms such as interaction with positive neighbourhood role models and individuals involved in positive practises. Whilst participation in extracurricular activities may be just one example of a difference in lifestyle, such an example may serve to identify how a greater impact of neighbourhood deprivation may arise amongst those with educated parents.

Alternatively, it may be argued that children of educated parents are susceptible to other factors which negatively influence educational outcomes, for example bullying, which could be influenced by characteristics associated with having an educated parent. With education possibly being correlated with risk aversion (Jianakoplos, 1998; Hersch, 1996), educated parents may control the child’s exposure to the neighbourhood, possibly restricting social networks and relationships with neighbours meaning that the child is isolated within school and vulnerable to bullying. Equally, with differential family characteristics, children of educated parents may fail to establish social or friendship groups with young people within the neighbourhood; Shin (2007) identifies a positive relationship between school performance and peer relationships.

2.6.3 Defining deprivation

There is no clear, accepted definition of neighbourhood deprivation when measuring deprivation by the IDACI score; initially neighbourhoods were defined as deprived if their scores were within the top 30% of the score distribution. It could be argued that a stricter definition of deprivation should be adopted thus observing those only in the more deprived neighbourhoods. Furthermore, the impact of neighbourhood deprivation may depend upon the scale of deprivation. For these reasons, a secondary definition of neighbourhood deprivation is introduced where neighbourhoods possessing an IDACI score within the top 20%, rather than 30% are observed.

The analysis will continue to follow the identical procedure as with the 30% definition, though the sample may change slightly when individuals who move in or out of deprived neighbourhoods within the observed time period are dropped.

Tables 2.10 and 2.11 present the relevant raw data for the 20% deprivation level analysis with the main results given in Table 2.12.

The overall neighbourhood effect for the 5A*-C GCSE outcome is now found to be insignificant. Hence, living within a neighbourhood that has an IDACI score ranked in the top 20% nationally, does not significantly influence the likelihood of obtaining 5 GCSEs A*-C relative to living in non-deprived neighbourhoods.

When observing the gold standard outcome, the neighbourhood effect is found to be smaller than that calculated when the 30% level definition of deprivation is adopted. This effect is indicating that individuals living in a deprived neighbourhood are 3.6 percentage points less likely to attain 5 GCSEs A*-C including English and maths relative to characteristically similar individuals living in a non-deprived neighbourhood. This is a significant effect but only at the 10% significance level.

When splitting the sample according to parental education and separately estimating the influence of neighbourhood deprivation upon GCSE outcomes, all individual estimates are insignificant for both those with educated and uneducated parents, equally for the five A*-C and the five A*-C including English and maths outcomes. Living within a neighbourhood with a deprivation rate in the top 20% according to IDACI scores therefore does not influence the likelihood of obtaining 5 GCSEs including or excluding English and mathematics, regardless of parental education.

These results differ substantially from those presented when a 30% deprivation rate was adopted; a higher scale of neighbourhood deprivation seems to present no negative influence upon outcomes. It may be expected that a greater degree of deprivation would more negatively impact upon individual outcomes; however, there are possible plausible explanations for these results.

Firstly, defining only neighbourhoods with a higher level of deprivation as deprived may capture largely neighbourhoods which are targeted by programmes or schemes that focus on the most deprived or very poor areas within England. These schemes may then assist in improving the GCSE prospects and achievements within very deprived neighbourhoods, thus offsetting the negative neighbourhood effect so that individuals in deprived areas are equally likely to obtain the observed GCSE outcomes as if they had lived in a non-deprived neighbourhood and therefore possibly been unaffected by such schemes. Moving from a 30% deprivation definition to a 20% definition would therefore mean that the outcomes of children within deprived neighbourhoods increase, due to the programmes targeting the very poor, to meet the outcomes of those in non-deprived areas. The identified neighbourhood effect estimates may then differ from those identified at the 30% level since the very poor neighbourhoods involved in programmes and schemes may account for a smaller proportion on the deprived neighbourhoods observed.

One example of such a scheme may be the Neighbourhood renewal fund which targeted the 88 most deprived authorities within England between 2001 and 2006, spending almost £1.9bn on the three broad areas of education, health and crime (Cowen and Wilton, 2008). Additionally, the SureStart children’s centres initiative, established in 1998, targeted the most 20% deprived neighbourhoods in England, providing health, social care and educational services in order to enhance the development of children; though introduced later than when LSYPE respondents were of the targeted age, such a programme may have influenced through their families’ interaction with the services for siblings. Other possible schemes may include teach first, who work only within schools where at least 50% of children are from the bottom 30% according to the IDACI, thus a lower neighbourhood deprivation level makes the school increasingly likely to be targeted. Though not an extensive list of all programmes, it is possible that a number of these targeted projects could influence the educational attainment of deprived neighbourhood residents, thus explaining an insignificant neighbourhood effect upon educational attainment when more highly deprived neighbourhoods are observed.

Alternatively, the identified insignificant influence of neighbourhoods with deprivation within the top 20% may be explained by the inclusion of previously defined deprived neighbourhoods in the newly defined non-deprived control group. It is possible that the negative neighbourhood influence is of the same magnitude for neighbourhoods within the top 20% and 30% deprived neighbourhoods alike. By observing the impact of the top 20% deprived neighbourhoods only, some neighbourhoods in the top 30% deprived thus become controls and enter the non-deprived group. The observed GCSE attainment within the deprived neighbourhoods therefore remains consistent with the attainment when observing all deprived neighbourhoods at the 30% level; the observed attainment within the non-deprived neighbourhoods however may be reduced relative to the main results since neighbourhoods inflicting negative effects are now included within the control group and these observations in the 20-30% deprivation range are probably likely to be the matched control observations when the PSM analysis is done. There is some evidence of this within the raw data; comparing Tables 2. 4 and 2.10 where the 30% and 20% deprivation level statistics are presented respectively, it is evident that the proportions obtaining the GCSE outcomes within deprived neighbourhoods are similar for the 20% and 30% levels. For the non-deprived neighbourhoods however, the alteration in definition from the top 30% to the top 20% deprived causes a fall in the proportions attaining both of the GCSE outcomes; for example, the change in definition causes the proportion of individuals attaining the gold standard within non-deprived neighbourhoods to fall from 55.7% to 52.7%; this is comparable to a change from 28.7% to 27% within deprived neighbourhoods.

2.6.4 Defining educated

As with the definition of a deprived neighbourhood, there is no clear consensus of what level of education should be deemed 'educated'. Initially, educated parents were defined as those with at least post-16 level education, however, broadly across empirical work, a definition of educated according to a degree is often adopted. For comparative purposes, this definition of an educated parent will be adopted. In doing so, the ratio of educated parents to uneducated parents becomes much smaller with 15.4% of the sample now deemed as having educated parents relative to 49.9% when adopting the post-16 definition (Table 2.6).

It is possible that the smaller sample sizes after matching cause the results to be less robust to a change in the matching procedure. For example, when matching of individuals with educated parents within deprived and non-deprived neighbourhoods, only 260 treated individuals could be matched. The nearest neighbour matching estimates will be discussed here, however it should be noted that differences arise in the estimates of the caliper matching procedure relative to the estimates of the NN matching.

The relevant raw data, given this change in definition, are given in Tables 2.6 and 2.13. Results are presented in Table 2.14.

Relative to the initial results, adopting a higher level of education definition of an educated parent causes a change in the overall neighbourhood effect since parental education is used to estimate the propensity to live within a deprived neighbourhood in order to match treatment and control individuals. This estimated neighbourhood effect is slightly higher than the initial results; ceteris paribus, those living in deprived neighbourhoods are 5.3 percentage points significantly less likely to attain 5 GCSEs A*-C relative to those in non-deprived neighbourhoods, and 8.7 percentage points significantly less likely to obtain the gold standard outcome. These estimated effects are highly significant and support the results of the main analysis to a certain extent. The results suggest that 21.4% of the raw gap in attainment of five GCSEs A*-C and 32% of the raw gap in the attainment of the gold standard between residents of deprived and non-deprived neighbourhoods may be explained by the neighbourhood effect.

As with the overall neighbourhood effect, when estimating the neighbourhood effects by parental degree education, the uneducated parent neighbourhood effect remains almost equivalent to the effect estimated in the main results. Results indicate that the likelihood of obtaining 5 GCSEs A*-C is insignificantly different for individuals with uneducated parents living in a deprived neighbourhood relative to those with uneducated parents, living in a non-deprived neighbourhood. However, those with uneducated parents in deprived neighbourhoods are around 6 percentage points less likely to attain the gold standard than similar individuals living in a non-deprived neighbourhood ceteris paribus. The neighbourhood effect thus explains 27.7% of the raw gap in the attainment of five GCSEs A*-C including English and maths between children living in deprived and non-derived neighbourhoods with uneducated parents.

Dissimilarities arise with the main results in the estimates of the neighbourhood effect for those with educated parents, now defined as parents with at least a degree level education. For individuals with educated parents, living in a deprived neighbourhood does not significantly influence the likelihood of attaining both 5 GCSEs A*-C and 5 GCSEs A*-

C including English and maths, relative to individuals in non-deprived neighbourhoods. There is therefore no evidence of a neighbourhood effect for individuals who have parents educated to degree level; these individuals are just as likely to obtain the GCSE outcomes of interest whilst living in a deprived area as if they had lived in a non-deprived neighbourhood. Whereas the main results indicated large and significant neighbourhood effects for individuals with educated parents, using the degree definition of educated dramatically changes these results.

One possible explanation for this dissimilarity could be that highly educated parents are more able to compensate for negative neighbourhood influences, thus, regardless of the neighbourhood deprivation rate, the child is equally likely to obtain the GCSE benchmarks. For example, educated parents may provide a higher quality of assistance with school work and exam preparation relative to parents with post-16 education only; this explanation does however contrast with the findings of McCulloch and Joshi (2001) who found that the home environment did not mediate the impact of neighbourhood deprivation. Alternatively, the young person may be more likely to aspire to attend university should their parent/s have done so, thus such aspirations may induce higher levels of effort in school which may influence attainment.

Another possible explanation for this insignificant effect of neighbourhood deprivation upon GCSE outcomes may be that children with educated parents to degree level are of higher ability, relative to children of parents with post-16 education. Consequently, in estimating the neighbourhood effect for those with degree level educated parents, neighbourhood effects do not appear to impact since these young people are easily able to obtain the 5 GCSEs A*-C and the gold standard outcomes regardless of their neighbourhood. Arguably, neighbourhood deprivation may still impact upon these individuals but not to such a great extent that they can no longer achieve these headline achievements. Such young people may suffer from neighbourhood effects in terms of their grades, for example achieving a B grade rather than A*, or obtaining seven GCSEs rather than nine for example. However, the outcome measures adopted would fail to capture a neighbourhood effect which operates in such a manner, since the focus of this study is to observe the impact on the predominant GCSE headline performance measures.

2.7 CONCLUSIONS

This chapter has investigated whether neighbourhood effects exist in determining educational outcomes at GCSE level, specifically observing the impact of neighbourhood deprivation upon the attainment of five GCSEs graded A* to C and five GCSEs A* to C including English and mathematics, also termed the gold standard. Using LSYPE data from 2003 to 2006, the differential effect of neighbourhood deprivation upon individuals with educated and uneducated parents was also examined in an attempt to answer the question: Are young people from uneducated families more susceptible to neighbourhood effects than the children of educated families?

The LSYPE provides a unique opportunity to investigate neighbourhood effects within England using recent data. The dataset provides a wealth of information on an individual level, supplying data on personal characteristics, attitudes, behaviours and achievements alongside, importantly, neighbourhood deprivation scores and respondent educational attainment information, such as GCSE attainments and prior test scores when the dataset is linked with the national pupil database.

In line with Harding (2003), the chapter adopts a propensity score matching procedure to estimate the impact of neighbourhood characteristics upon individual outcomes. The overall neighbourhood effect is estimated using PSM techniques and subsequently the neighbourhood effects by parental education, by separating the sample accordingly. PSM methods are advantageous in the estimation of neighbourhood effects since the matching procedure alleviates the main issues surrounding the measurement of neighbourhood effects namely the issues of a selection bias, causality and the evaluation problem.

The main analysis investigates the influence of neighbourhood deprivation upon educational attainment when defining deprived neighbourhoods, initially, as those within the top 30% deprived, according to IDACI scores. The primary analysis involves estimating the neighbourhood effect overall for the full sample. Ceteris paribus, neighbourhood deprivation negatively influences the probability of obtaining five GCSEs A*-C and five A*-C including GCSEs in English and mathematics. Individuals living in deprived neighbourhoods for at least the three years, are around 4 percentage points less likely to obtain five GCSEs A*-C, relative to individuals living in non-deprived neighbourhoods and are around 6 percentage points less likely to obtain the gold standard GCSE outcome. The neighbourhood effect may therefore explain 16.1% of the raw gap in the attainment of five GCSEs A*-C, and 22.2% of the raw gap in the attainment of the gold standard between deprived and non-deprived residents. These results reflect a common finding throughout whereby of the significant neighbourhood effects identified, neighbourhood deprivation has a greater influence on the attainment of the 5 A*-C including English and maths than the standard 5 A*-C outcome.

The overall neighbourhood effect is additionally calculated when neighbourhood deprivation and parental education are redefined. When investigating whether a stricter definition of a deprived neighbourhood influences the neighbourhood effect estimates, by adopting a definition of deprived as neighbourhoods with scores within the top two deciles of the IDACI, findings suggest a smaller neighbourhood effect with a significant impact of neighbourhood deprivation upon the gold standard outcome only.

In the analysis of neighbourhood effects by parental education, the impact of neighbourhood deprivation upon GCSE outcomes is estimated separately for individuals with educated and uneducated parents; the difference between these neighbourhood effects is then calculated. From the main analysis, negative and significant neighbourhood effects are identified for individuals with educated parents with at least post-16 education; ceteris paribus, individuals with educated parents living in deprived neighbourhoods are around 7 percentage points less likely to obtain 5 GCSEs A*-C, and around 12 percentage points less likely to gain the gold standard, relative to characteristically similar individuals with educated parents from non-deprived neighbourhoods, based upon nearest neighbour matching estimations. Neighbourhood effects therefore seem to explain 35.9% of the raw gap in the attainment of five GCSEs A*-C and 51.7% of the raw gap in the attainment of five GCSEs A*-C including English and maths between deprived and non-deprived residents with educated parents.

Neighbourhood deprivation is found to influence individuals with uneducated parents to a lesser extent; whilst insignificantly impacting upon the attainment of five GCSEs A*-C, the likelihood of obtaining the gold standard is reduced by around 6 percentage points by living in a deprived neighbourhood for young people with uneducated parents. The estimated neighbourhood effect is significantly larger for individuals with educated parents signalling that the penalty associated with neighbourhood deprivation imposed upon the educational attainment of residents is greater for individuals with educated parents who would benefit to a greater extent by living in a non-deprived neighbourhood, relative to individuals of uneducated parents. Speculation and proposed explanations for this identified differential relate to neighbour and peer ability, aspirations and peer group choices.

Neighbourhood effects are additionally estimated when adopting a 20% definition of deprived neighbourhoods and when defining ‘educated’ as degree level education rather than post-16 level as in the main analysis. When the stricter definition of deprivation is adopted, all neighbourhood effects both upon the attainment of five GCSEs A*-C and upon the gold standard are insignificant; this is so for both individuals with educated and uneducated parents. A number of explanations are postulated for this identified effect including reasons associated with poor neighbourhood targeting programmes, which may reduce the negative neighbourhood effect, and alternatively, underlying individual characteristics that may predetermine ‘bad’ outcomes. Defining parents as educated when holding a degree and subsequently estimating the neighbourhood effect gives similar results as in the main analysis for individuals with uneducated parents. However, variation from the estimates within the main analysis is evident within the estimates of the neighbourhood effect for individuals with educated parents; ceteris paribus, the impact of neighbourhood deprivation upon the GCSE attainment of individuals with parents educated to at least degree level is insignificantly different from zero when considering both five A*-C and gold standard GCSE outcomes. It is suggested that neighbourhood deprivation may remain to impede upon education but this is uncaptured within this analysis which focuses on broad headline measures.

To summarise, the main analysis within this chapter reveals an interesting finding; neighbourhood effects are found to be negative and significant, thus contrasting with the findings of other neighbourhood effects studies (Gibbons, 2012; Weinhardt, 2013; Sanbonmatsu, 2006; Lindahl, 2008; McCulloch and Joshi, 2001). A possible explanation for the differential results both between this study and other neighbourhood effects papers and amongst the neighbourhood literature is the variation in methods across studies; there is not a clear single method which has been adopted or identified as being the most suitable in estimating the impact of neighbourhoods upon outcomes such as education. In addition the data adopted, the definition of a neighbourhood, the deprivation measure or index and the outcome of interest varies between studies thus explaining the range of findings within the neighbourhood effects literature.

In addition to the identified overall negative neighbourhood effect, this study finds that the GCSE outcomes of individuals with educated parents are identified as being diminished to a greater extent by living in a deprived neighbourhood, relative to individuals with uneducated parents. Whilst presenting an alternative approach to measuring the impact of neighbourhood deprivation upon educational attainment, this chapter additionally presents further analysis of neighbourhood effects by identifying the family background characteristics of individuals who may be more susceptible to the negative influences. This may be important for policy since the results indicate that targeting children based upon their socio-economic status may fail to aid those with educated parents whose educational attainment may suffer due to deprived surroundings. It is not only children from deprived and uneducated families who fail to reach their potential within deprived neighbourhoods, it is more so the children of educated parents whose may potentially be more able but suffer educational losses due to the neighbourhood in which they live. The findings of this chapter highlight the importance of the definition of neighbourhood deprivation adopted within the measurement of neighbourhood effects.

It is important to acknowledge the shortcomings of this study however. This empirical analysis could be improved by adopting panel data which tracks or informs of previous residence, prior to year 9 of the schooling system as in the LSYPE, allowing for individuals with longer term exposure to neighbourhood deprivation to be separated from individuals living in deprived neighbourhoods for the minimum term, i.e. the three years observed within this study. This would therefore allow for the analysis of long and short term neighbourhood effects to be identified, as the length of exposure has been signalled to influence outcomes within the literature (Chetty et al. 2016). In addition, it may be of interest to consider continuous outcome variables such as the number of GCSEs attained. This may, for example, more precisely capture whether neighbourhoods impede upon the attainments of those with parents educated to degree level. The finding of insignificant neighbourhood effects for children of degree level educated parents may of course be due to data restrictions and sample size since only a small number of individuals with degree educated parents within deprived neighbourhoods were successfully matched; a larger sample size and dataset may therefore have benefitted this part of the analysis within this chapter.

Table 2.1 Determinants of Attrition

	Attrition between waves 1-3 **(N'hood deprivation control)**	**Attrition between waves 1-3** **(Full controls)**
Deprived neighbourhood 30%	0.075 (0.00)	0.019 (0.024)
Parental education post-16	-	-0.026 (0.00)
Household employment	-	-0.014 (0.203)
Professional parent	-	-0.006 (0.518)
KS2 ability	-	0.003 (0.681)
KS2 ability squared	-	-0.000 (0.204)
Born in UK * white	-	-0.041 (0.00)
Household deprivation	-	0.018 (0.114)
Parental interest: homework*parents evening*intentions for educ.	-	-0.023 (0.004)
School record A*-C	-	0.024 (0.038)
School interaction: A*-C record,, Class size abv av. & Mainstream school	-	-0.024 (0.01)
N	*15,767*	*10,424*

P values in parenthesis

Table 2.2 Characteristic controls descriptive statistics

Variable	variable type	Mean	Standard deviation
Household employed	Binary	0.775	0.418
Parental education (post-16 educated)	Binary	0.483	0.362
Professional parent	Binary	0.312	0.463
KS2 ability	Continuous	27.27	3.920
KS2 ability squared	Continuous	757.39	202.975
Born in UK & white	Binary	0.698	0.459
Household deprivation	Binary	0.196	0.397
Parental interest: homework*parents evening*intentions for educ.	Binary	0.441	0.497
School record A*-C	Binary	0.826	0.379
School interaction: A*-C record,, Class size abv av. & Mainstream school	Binary	0.641	0.480

Table 2.3 Balancing checks

	30% full sample	Educated parents	Uneducated parents
Pseudo R2	0.001	0.002	0.001
Hotelling	0.253	0.634	0.829
Standardized bias (%)	1.844	3.297	1.749
Absolute bias (highest)	6	8	3
T-stat	All insignificant at 1% level except 1 covariate: P-value=0.007	All insignificant at 5% level	All insignificant at 1% level

Table 2.4 Proportion of individuals within deprived/ non-deprived neighbourhoods attaining GCSE outcomes (30% deprivation)

	Deprived neighbourhood	Non-deprived neighbourhood
5 GCSEs A*-C	41.9%	66.7%
5 GCSEs A*-C inc. English and maths	28.7%	55.7%

Table 2.5 Proportion on individuals within deprived/ non-deprived neighbourhoods attaining GCSE outcomes, by parental education (30% deprivation)

Attainment:		Deprived neighbourhood	Non-deprived neighbourhood
5 A*-C	**Educated parents Post-16**	56.1%	76.7%
	Uneducated parents Post-16	35.7%	52.2%
5 A*-C inc. English and maths	**Educated parents Post-16**	42.5%	66.3%
	Uneducated parents Post-16	22.8%	40.3%

Table 2.6 Proportion of sample with educated and uneducated parents (30% deprivation)

	Post-16 education	Degree
% with educated parents	49.9%	15.4%

Table 2.7 Proportion of educated and uneducated parents within deprived and non-deprived neighbourhoods (30% deprivation)

	Deprived neighbourhood	Non-deprived neighbourhood
% with educated parents Post-16 education	30.2%	59.0%
% with uneducated parents Post-16 education	69.8%	41.0%
Total	100%	100%

Table 2.8 Location of the educated parents by deprivation

	Deprived neighbourhood	Non-deprived neighbourhood	Total
% with educated parents post-16	19.1%	80.9%	100%

Table 2.9 Propensity score matching: 30% deprivation Post-16 education definition

Educated: Post-16 Education / Deprivation: Top 30% deprived IDACI									
	Outcome: 5 A-C*				*Outcome: 5 A*-C including Eng & Mat.(gold standard)*				
	(1) **Propensity score Nearest neighbour**	(2) **Difference: uneducated and educated**	(3) **Propensity score Caliper matching**	(4) **Difference: uneducated and educated**	(5) **Propensity score Nearest neighbour**	(6) **Difference: uneducated and educated**	(7) **Propensity score Caliper matching**	(8) **Difference: uneducated and educated**	**N (Treated)**
Neighbourhood effect (full sample)	-0.040* (0.018)		-0.041* (0.018)		-0.060*** (0.016)		-0.061*** (0.016)		3352
Neighbourhood effect educated parents	-0.074** (0.027)	-0.057 (0.035)	-0.079** (0.027)	-0.063 (0.034)	-0.123*** (0.028)	-0.067* (0.033)	-0.128*** (0.026)	-0.071** (0.032)	1309
Neighbourhood effect uneducated parents	-0.017 (0.022)		-0.017 (0.022)		-0.057** (0.019)		-0.057** (0.019)		2512

Significance: *** 1% level **5% level *10% level

Table 2.10 Proportion of individuals within deprived/ non-deprived neighbourhoods attaining observed GCSE outcomes (20% deprivation)

	Deprived Neighbourhood	Non-deprived neighbourhood
5 GCSEs A*-C	40.7%	63.8%
5 GCSEs A*-C inc. English and maths	27%	52.7%

Table 2.11 Proportion of individuals within deprived/ non-deprived neighbourhoods attaining GCSE outcomes, by parental education (20% deprivation)

Attainment:		Deprived neighbourhood	Non-deprived neighbourhood
5 A*-C	**Educated parents Post-16**	53.9%	75.1%
	Uneducated parents Post-16	35.7%	49.4%
5 A*-C inc. English and maths	**Educated parents Post-16**	39.4%	64.6%
	Uneducated parents Post-16	22.3%	37.5%

Table 2.12 Propensity score matching: 20% deprivation Post-16 education definition

Educated: Post-16 Education / Deprivation: Top 20% deprived IDACI									
	Outcome: 5 A-C*				*Outcome: 5 A*-C including Eng & Mat.*				
	(1) **Propensity score Nearest neighbour**	(2) **Difference: uneducated and educated**	(3) **Propensity score Caliper matching**	(4) **Difference: uneducated and educated**	(5) **Propensity score Nearest neighbour**	(6) **Difference: uneducated and educated**	(7) **Propensity score Caliper matching**	(8) **Difference: uneducated and educated**	**N (Treated)**
Neighbourhood effect (full sample)	-0.014 (0.017)		-0.014 (0.017)		-0.036* (0.017)		-0.036* (0.017)		2507
Neighbourhood effect educated parents	-0.021 (0.030)	-0.036 (0.038)	-0.021 (0.029)	-0.038 (0.037)	-0.056 (0.033)	-0.033 (0.040)	-0.060 (0.031)	-0.038 (0.039)	662
Neighbourhood effect uneducated parents	0.015 (0.022)		0.016 (0.022)		-0.023 (0.023)		-0.022 (0.024)		1845

Significance: *** 1% level **5% level *10% level

Table 2.13 Proportion of individuals within deprived/ non-deprived neighbourhoods attaining GCSE outcomes, by parental education (degree level educated definition)

Attainment:		Deprived neighbourhood	Non-deprived neighbourhood
5 A*-C	Educated parents degree level	73.8%	89.3%
	Uneducated parents below degree level	39.6%	61.8%
5 A*-C inc. English and maths	Educated parents degree level	66.1%	82.2%
	Uneducated parents below degree level	26.1%	49.2%

Table 2.14 Propensity score matching: 30% deprivation, degree education definition

Educated: Degree Education / Deprivation: Top 30% deprived IDACI									
	Outcome: 5 A-C*				*Outcome: 5 A*-C including Eng & Mat.*				
	(1) **Propensity score Nearest neighbour**	(2) **Difference: uneducated and educated**	(3) **Propensity score Caliper matching**	(4) **Difference: uneducated and educated**	(5) **Propensity score Nearest neighbour**	(6) **Difference: uneducated and educated**	(7) **Propensity score Caliper matching**	(8) **Difference: uneducated and educated**	**N (Treated)**
Neighbourhood effect (full sample)	-0.053*** (0.015)		-0.053*** (0.015)		-0.087*** (0.017)		-0.087*** (0.017)		3352
Neighbourhood effect educated parents	-0.023 (0.043)	0.0045 (0.047)	-0.045 (0.038)	-0.017 (0.042)	-0.035 (0.050)	-0.030 (0.053)	-0.045 (0.049)	0.019 (0.052)	260
Neighbourhood effect uneducated parents	-0.028 (0.020)		-0.028 (0.019)		-0.064*** (0.019)		-0.064*** (0.018)		3282

Significance: *** 1% level **5% level *10% level

CHAPTER 3 : SETTING IN MATHS AT PRIMARY SCHOOL AND CHILD BEHAVIOUR: EVIDENCE FROM THE MILLENNIUM COHORT STUDY

3.1 INTRODUCTION

Grouping children by ability is a practise that has been adopted widely in the UK both in primary and secondary schools. Over the past 80 years, a number of grouping strategies have been adopted within schools as a result of the Hadow report in 1931; one such method of ability grouping is streaming, which refers to the 'division of all pupils in a year group into classes hierarchically structured according to a measure or judgement of 'overall' academic ability' (Campbell, 2013). Streaming peaked in popularity through the 1940s to 1950s and remains an option for ability grouping, though the incidence of streaming is rather small (Hallam et al., 2003). Alternative grouping practises include setting, which involves dividing pupils within a year group into classes according to measured or perceived ability for the teaching of a given subject, in addition to within-class ability grouping. Within-class ability grouping involves dividing a class into sub-groups, based on measured or perceived ability, for general teaching purposes or for teaching in a specific subject (Campbell, 2013). Setting and within-class ability grouping strategies are much more common within UK schools than streaming and have more recently been recommended by policy. The 1997 Labour government widely encouraged setting within schools, leading to a growth in the incidence of setting; current policy however provides little guidance on ability grouping practises though there is evidence in datasets such as the Millennium Cohort Study (MCS) that ability grouping practises continued to be implemented within schools.

With recommendations dating back to the 1960s, setting in particular is an interesting form of ability grouping since the longevity of setting within UK schools suggests a practise which schools, teachers and possibly children have gained experience and understanding of, whilst the current prevalence within schools means that the impact of setting is still relevant today.

The policy of class setting involves dividing children of the same year group into classes for a specific subject based upon ability; such a policy does not go without opposition. Supporters of setting argue that teaching may be more efficient since teachers may target the ability level and the needs of students; this may reduce the likelihood of disengagement of lower ability children relative to when teachers must provide lessons for heterogeneous classes with children of all levels of ability (OFSTED, 2000). Those in opposition of the policy argue that peer effects would benefit the lower ability pupils when taught with higher ability children in mixed classes; in addition, it is also argued that the act of setting may be demeaning for children whose confidence and motivation are damaged by setting (Kutnick et al., 2005). It is also argued that children may act out the roles assigned to them, thus behaviour and attainment may diminish as a result of being placed in a lower set.

The impact of setting alongside ability grouping more generally, has been examined predominantly considering attainment and achievement outcomes; though some research has attempted to explore the impact of ability grouping on alternative outcomes, such as self-concepts and self-esteem (Abadzi, 1985; Ireson and Hallam, 2009), few studies have examined behavioural outcomes. In addition, little research has been undertaken within the field of economics, whilst few studies have adopted advanced methodological approaches or techniques to overcome some of the surrounding econometric problems, providing the basis of this chapter. The research that has been undertaken so far fails to reach a clear consensus; whilst some studies find little evidence of setting influencing outcomes (Barker-Lunn 1970; Kulik and Kulik 1982, Ireson and Hallam 2005), a number of studies have identified that whilst high ability pupils benefit from being grouped by ability, mixed ability classes advantage lower ability pupils (Hallam and Parsons 2014; Ireson 1999a: Suknandan and Lee 1998; Slavin 1988).

By observing the non-cognitive outcomes of children, specifically child behaviour, this chapter attempts to aid the growth of our understanding of the determinants of non-cognitive outcomes. Within the economics field, cognitive outcomes have received much greater attention within existing studies than non-cognitive outcomes, despite the fundamental role that non-cognitive development plays in child progress and individual life chances (Vignoles and Meschi, 2010); child behaviour in particular has been identified as a significant determinant of schooling outcomes (Kirstoffen and Smith, 2013). By adopting econometric techniques to analyse the impact of setting on behaviour, this chapter will contribute to the literature which focuses on the non-cognitive development of children alongside the ability grouping research.

This chapter seeks to identify the impact of setting children by ability in mathematics on behaviour, when children are at primary school. Data from the UK Millennium Cohort Study (MCS) is adopted within this chapter, specifically from waves 4 and 5 of the survey when children are aged 7 and 11 respectively. This is a suitable dataset since it provides information on whether the child experiences setting in each wave and the level of set placement. Setting in maths specifically is the focus of the chapter since this is the subject for which ability setting is more prominent in schools around the UK (Hallam et al., 2003), but also in the adopted dataset. In addition, since maths is a core subject, children spend a large proportion of their school week in their allocated class for this subject. Child behaviour is measured by the Strengths and Difficulties Questionnaire (SDQ) which is a behavioural screening questionnaire. This chapter observes three outcome measures of behaviour provided by the SDQ; the total difficulties score, internalising behaviour and externalising behaviour. Both the responses of the parent and the teacher are observed, allowing for a comparison between home and school behaviours.

The questions that this chapter would like to address are firstly; does setting influence behaviour? To investigate this, Ordinary Least Squares (OLS) and Fixed effects (FE) estimation approaches are adopted. Due to the potential issue of unobserved heterogeneity due to underlying unobserved characteristics of the child potentially influencing behaviour, the FE model is favoured.

Secondly, the chapter investigates if behaviour is influenced by the level of the set in which pupils are placed. Specifically, the impact upon behaviour is investigated for the children who are placed in the lowest set for maths since these are the children who are most often the centre of the setting debate. In examining the impact of being placed in a low maths set, the potential problem of endogeneity must be addressed, this arises since it is likely that the behaviour of a child influences the set in which they are placed. Whilst OLS is initially adopted, an Instrumental Variable (IV) approach is additionally taken to overcome this issue.

One additional concern of this chapter is to identify whether a gender differential exists in both the impact of setting and set placement in maths. Since evidence suggests that the behaviour of girls and boys may not respond in a similar manner (McNeish and Scott, 2014; Leadbeateret al. 1999), the chapter seeks to identify whether this is also true when observing the impact of the school policy of setting. Little research has been undertaken in this area, thus this investigation of gender differentials will add to the existing literature. The chapter will also contribute to the setting, and more generally, the ability grouping research, through the investigation of the impact of class setting upon behaviour, as measured by SDQ scores, which to my knowledge has not been examined previously. Furthermore, the methods adopted within this chapter present a more econometric based investigation into the influence of setting relative to the existing literature; the methods adopted within this chapter attempt to overcome the issues of unobserved heterogeneity and endogeneity that are faced by researchers when estimating the impact of setting. One final contribution that the chapter will make will be to add to the research based upon setting and ability grouping in primary schools since few studies consider the impact at this level of schooling or age; existing literature predominantly concerns itself with ability grouping in secondary schools in the UK or equally high school in the US.

The chapter will be structured as follows: the background of ability grouping and setting will be discussed in section 2 where definitions of each of the practises will also be provided. A literature review will be provided in section 3. The data and methodologies will be discussed within sections 4 and 5 respectively whilst section 6 will provide descriptive statistics and an examination of the results. The chapter will close with a summary of the chapter and the results in the conclusion located in section 7.

3.2 BACKGROUND

Within the UK, there is a history of ability-grouping in schools; the Hadow report (1931) recommended that primary schools that were large enough should adopt streaming measures to group children by ability. Streaming was a popular policy within primary schools over the following decades until the 1960s when research suggested that there were negative consequences of streaming (Jackson, 1964). The Plowden report (1967) subsequently promoted 'unstreaming' and advocated other forms of ability grouping such as within-class ability grouping and setting for certain subjects, though the introduction of the national curriculum from the late 1980s led to an increase in whole-group teaching in Primary schools (Pollard and Triggs, 2000).

Following a report from the Department for Education (1993), class setting specifically as a form of ability grouping was promoted and encouraged within all primary schools.

This was later reinforced by the newly elected Labour government in 1997 who widely encouraged ability grouping and setting in particular, and emphasised the need to raise standards in the UK educational system. The 2005 white paper suggests that “Grouping students can help to build motivation, social skills and independence; and most importantly can raise standards because pupils are better engaged in their own learning. We have encouraged schools to use setting since 1997” (House of Commons, 2005). The paper also states that whilst schools will be encouraged to group by ability, the decision whether to adopt this policy or not lies with the individual school.

Class setting is a school policy that seems to have gained prevalence amongst primary schools within the UK during the Labour period; Hallam et al. (2003) found that 60% of junior schools and 50% of infant and junior schools set students for at least one subject. Schools that set students were most likely to use class setting for mathematics whilst setting was more prevalent for the older year groups within the school, with most schools setting in years 5 and 6 only when children are aged 9-10 and 10-11 respectively.

The 2010 coalition government provided little backing or objection to any form of ability grouping, including setting, with little mention within policy. Following reports of a rumoured policy to implement compulsory setting by conservative ministers, the government made clear that it did not advocate setting and stated that schools were left to decide on the organisation of their teaching (TES, 2013). Though the coalition government implemented little policy change in relation to class setting or any other form of ability grouping within both primary and secondary schools, it is unclear whether the conservative government will do so; this is despite the reported considerations of the conservative government to implement compulsory class setting within schools (BBC, 2014). In spite of the indeterminate policy backing, the practise continues to be established within many schools.

3.3 LITERATURE REVIEW

3.3.1 Support and opposition for class setting: Theoretical discussion

Critics of class setting argue that peer group effects have an important influence within mixed ability groups; high achievers and highly driven students are able to motivate and stimulate the students within the class, thus potentially increasing attainment. By separating these higher ability students, it is argued that lower ability students in particular are harmed. Eilam and Finegold (1992) argue that separating children by ability deprives the low attaining students of academic role models and limits peer support, thus affecting the motivation, attitudes and behaviour of low ability pupils.

One argument against setting, or other forms of ability grouping, is the potential adverse effect for lower ability pupils who may be demotivated by being placed in lower sets (Kutnick et al. 2005). The labelling of pupils through grouping is likely to affect self-perceptions alongside behaviour through being either implicitly or explicitly informed of their ability level relative to other classmates and gaining knowledge of the level of their placement; in turn, pupils may behave and perform at the correspondingly high or low level (Campbell, 2013).

In addition, Kutnick et al.(2005) argue that lower ability grouped pupils may develop anti-school attitudes and may be de-motivated by their placement in lower sets which may cause a slower rate of progression. The progress of higher placed pupils may on the other hand be influenced by positive attitudes and expectations derived from higher set placement. Relatedly, the teaching environment may differ between lower and higher ability groups; Oakes (1985) identified that when children are grouped by ability, within the higher ability groups peers were supportive of one another whereas in lower ability classes hostility and anger characterised peer interactions.

One further concern with ability grouping is the tendency for pupils to be segregated by factors such as ethnicity and social class since attainment and ability tend to be stratified along these factors. Classes of higher ability pupils have a propensity to contain a lower proportion of pupils from ethnic backgrounds and from a lower social class, thus it is argued that the segregation of children involved in setting aids a widening of the social gap (Gamoran, 2002). In addition, lower sets are also found to contain more younger children in the year group, born in the spring or summer (Ireon et al. 1999), which may exacerbate the already existent gap between the spring/ summer born and the autumn / winter born children (Campbell, 2013).

Conflicting with these views, supporters of class setting argue that grouping by ability provides a more efficient basis for teaching since classes may be tailored to the ability level of the group (Gamoran, 2002). OFSTED (2000) states that in schools using setting, teachers found teaching easier since being able to target a narrower attainment range. The pace of teaching may be adapted to the group, thus allowing for lower ability students to engage in lessons whilst higher ability pupils are not held back. In addition, an environment in which students may progress at a comfortable pace without feeling pressured by the higher capabilities of highly able pupils may be generated.

Related to this, lower ability students may be less likely to detract from the class when they are able to understand and engage in the lesson and the appropriate level material, since in mixed ability classes lower ability pupils may find lessons ‘meaningless’ when they are unable to engage in the teaching and lesson and are therefore likely to either do little work or ‘act up’ (House of Commons, 2011).

3.3.2 Existing research discussion

The topic of ability grouping including class setting is a subject that is often debated within the educational and psychological fields; this applies when considering the effect upon outcomes such as achievement, self-concepts and attitudes to schooling. The impact of setting is likely to encompass two broad effects which are unlikely to be independent; firstly, there is likely to be a peer effect involved in sorting children according to ability. The classroom peers within the sets may influence the teaching environment through factors such as behaviour, motivation or simply interest and attentiveness, thus potentially influencing the learning experience alongside attitudes, behaviour and self-concepts of other pupils within the class. It is also likely that set placement may influence motivation and self-perceptions when pupils are implicitly or explicitly informed of their relative ability level. Secondly, ability grouping practises such as setting are likely to involve a change in teaching strategy or approach; teachers are able to narrow their instruction according to the ability of the class and subsequently provide exercises, resources and facilitation in line with the class ability. This in itself is likely to influence the child’s learning but this may also influence other factors such as behaviour.

While a number of papers within the education and psychology literature attempt to observe the impact of ability grouping including class setting, few seek to identify the impact upon behavioural issues since existing studies primarily consider the impact upon student achievement and attainment in terms of test scores. However, the attainment and behaviour research may be interrelated since it is possible that behaviour is impacted by setting which in turn effects attainment. Kristoffersen and Smith (2013) find that behavioural problems, as identified by SDQ measures, impact upon school outcomes; the influence of behaviour on outcomes may depend on gender and the type of behavioural problem, such as abnormal externalising behaviour. These considerations will be made within this chapter.

In addition, few studies have adopted econometric techniques to overcome the potential issues of unobserved heterogeneity and endogeneity when analysing the impact of setting, with many papers adopting qualitative or simply comparative methodologies.

Due to the literature on setting in primary schools being rather scarce, the literature review will discuss the impact of other ability grouping practises in addition to the impact of grouping in a secondary school setting. Though the outcome measures vary, the methods and considerations made within these studies remain relevant to this chapter. The recent literature which examines the influence of ability grouping on attainment will be summarised before moving onto the evidence of the alternative effects of class setting on factors such as behaviour and self-concepts, which are more directly related to the focus of this chapter.

3.3.3 Class setting and attainment

In 1998, OFSTED attempted to identify the prevalence of setting by ability within primary schools and the effects of setting by analysing setting from the survey data of 900 randomly selected schools. The report indicated a high prevalence of setting by ability in mathematics with years 5 and 6 being the most likely school years that schools adopted a setting policy. OFSTED reported impressive gains in the national tests of ‘setted’ subjects and subsequently advised that primary schools adopt setting by ability as a means to raise standards (OFSTED, 1998).

However, many studies have failed to identify a positive effect of ability grouping upon achievement relative to mixed ability groups overall. In the case of UK primary schools and setting in mathematics specifically, Whitburn (2001) provides an investigation using data on 1200 children in a single London borough to evaluate whether primary school children set for mathematics attain greater test scores in a routine short maths test relating to the previous term’s curriculum material. The comparative study examines the change in test results of set children relative to mixed-ability taught children by observing KS2 children in year two and year three. The study also investigates the impact of setting upon the variation in attainment within the class; the findings indicate that children of all attainment levels achieve greater results when taught in mixed ability classes whilst surprisingly the range of attainment within the class was smaller within mixed ability classes. The findings provide little evidence of a positive role of setting in determining child attainment at primary school. Mixed ability teaching is recommended since the results suggest equitable benefits for pupils.

More recently, a paper by Hallam and Parsons (2014) investigates the impact of streaming on the academic progress of children in year 2 of UK primary schools. Adopting data from the MCS, the study considers whether there is a heterogeneous effect of streaming upon children of different ability levels. Taking account of individual, family and school characteristics and controlling for attainment at age 5, the study adopts multiple regression analysis to evaluate impact of streaming on the assessment results of streamed children relative to children who were not streamed. Assessment results in maths, reading and an overall indicator of reading, writing, maths and science were observed. Findings suggest that students placed in a top stream benefit from streaming in terms of attainment relative to non-streamed children; however, those in the lower and middle streams were disadvantaged in terms of reading and overall results relative to the non-streamed. The lower streamed pupils in particular performed significantly worse in maths than the non-streamed. The findings of Hallam and Parson provide an interesting conclusion that the impact of setting may not be a homogenous effect for all ability levels; whilst grouping may be advantageous for higher ability, it is detrimental to the lower ability groups. It may therefore be more beneficial for lower ability pupils to be taught in mixed ability classes though this would be at the cost of higher ability pupils.

The findings of Hallam and Parsons (2014) are similar to those of Ireson (1999a) who presents the results from the Ability Grouping in Secondary Schools Project, which began in 1997 following the increased policy interest in class setting. The project attempted to examine the relationship between ability grouping on the attainment of year 9 pupils in English, maths and science, alongside non-academic outcomes such as attitudes to schooling and self-esteem. The data includes 6,000 students from 45 mixed secondary schools across London, Southern counties of England, East Anglia and South Yorkshire. Schools within the sample may be split into three categories according to ability grouping levels: mixed ability setting where setting occurred in no more than two year 9 subjects, partially set schools where setting was adopted in a maximum of two year 7 subjects and four year 9 subjects or set schools where streaming, banding or setting was adopted in at least four subjects from year 7. A multilevel analysis methodology was adopted to take account of data variation at the pupil level and school level simultaneously. Implementing a fixed effects methodology alongside a random effects modelling approach, the study assessed the impact of setting for each curriculum subject individually. The results provide little evidence that progress in English and science was associated with setting; however, in mathematics, the lower attaining students at KS2 made more progress in mixed ability grouping schools whilst the higher attaining pupils benefitted in schools with more setting. The higher attaining pupils gained from setting to a greater extent than the lower attaining pupils gained from mixed ability classes.

One paper of particular interest to this chapter is Betts and Shkolnik (2000), due to the similarities in methodologies. This paper examines the impact of setting on student achievement growth in mathematics and the allocation of resources in US secondary schools, using data from the Longitudinal Study of American Youth (LSAY). Two cohorts of pupils are observed with the first cohort observed between grades 7 and 9 and the second cohort between grades 10 and 12[6]. The study firstly adopts a typical education production function to estimate the net effect of ability grouping. The study goes on to examine whether a differential effect of setting exists dependent on ability by estimating an additional model which estimates separate Ordinary Least Squares (OLS) equations for each group of students (high ability, mid ability, low ability and heterogeneously grouped students). Little evidence of an effect of ability grouping on math achievement growth is identified while no differential effects of grouping are identified upon the various ability levels. As a test of robustness, an IV approach is adopted within the estimation of the differential effects of setting by ability levels; the set placement of pupils is instrumented using three instruments; firstly, the percentage of black pupils in the school, secondly, the percentage of students receiving full federal lunch assistance in the school and finally, the pupil's test score relative to the average for their grade. The results provide evidence that ability grouping influences math progress within this robustness analysis where the lower groups are found to be unaffected by grouping, the mid groups and found to be negatively influenced and the top group is identified as benefitting from ability grouping practises. These identified effects are however reasonably small and are suggested to provide limited evidence only. When including achievement quartiles within the regression rather than mean class ability controls, grouping is found to have no effect on achievement.

[6] Grade 7 aged 12-13
Grade 9 aged 14-15
Grade 10 aged 15-16
Grade 12 aged 17-18

Other studies have diverged from this frequent finding that ability grouping benefits the higher ability pupils whilst disadvantaging the lower ability; Ireson et al.(2005) attempt to identify the impact of setting upon GCSE grades in English, maths and science, using the same data as Ireson (1999a)[7]. The study adopts a multi-level methodology to estimate the impact of setting both between leaving primary school and KS3 and to GCSE examinations. A value-added approach is therefore essentially adopted in estimating the impact of setting on each of the core subjects individually. The initial findings indicate little influence of years of setting upon GCSE attainment in any of the core subjects. The impact of the level of the set of placement is additionally analysed using simple regression analysis; the results suggest that there is little differential impact of setting upon high and low attaining pupils in English and maths, though in science there is a small negative impact on high attaining students and a positive effect upon low attaining students. This study identifies an almost opposite effect to the prior papers discussed when observing the impact of setting in science though other findings agree somewhat with the earlier papers that identify little influence of setting on attainment.

One further study from the US looks at the impact of tracking, which is the US equivalent of streaming, upon maths and reading scores. The study by Collins and Gan (2013) utilised student-level data encompassing 9,325 children from across Dallas; this includes data from two school years when students are initially in the third grade, when aged 8-9 in 2003-4, and in the fourth grade when aged 9-10 in 2004-5. The study looks at the extent to which the degree of sorting within a class impacts upon student achievement by exploiting the variation in sorting between school years. This approach involves the construction of a sorting index, indicating how "sorted" a class is; in addition, the effect of sorting upon different types of students is considered. The paper adopts an IV methodology to overcome the possible endogeneity problem that arises in measuring the effect of ability grouping due to unobserved characteristics determining both the student's placement and their achievement; one example of this may be the student's behaviour. The sorting in one school year is instrumented by the sorting index from another school year at the same school, since the ability grouping policies through the school should be related to the particular sorting within a year group yet the sorting in another year group at the school should not determine or influence the students' attainment. The findings suggest that students' reading and maths scores benefit from more homogenous classes. The study also finds that the influence of sorting on high scoring students is

[7]The data tracked 6,000 students from within 45 mixed secondary comprehensive schools across England

insignificantly different to the impact on low scoring students, suggesting that both high and low ability benefit from ability grouping. This finding deviates from the results of other papers; though the discussed literature provides rather inconsistent evidence, only this study indicates a beneficial effect of tracking for both high and low ability pupils. What may be interesting is the use of advanced methodological techniques used to address endogeneity in this study relative to other papers within this research area.

The existing research on the impact of ability grouping on attainment provides little evidence of a strong relationship. A number of papers identify that ability grouping may benefit higher ability pupils while being detrimental to lower ability pupils; this finding is consistent with earlier studies such as Suknandan and Lee (1998) and Slavin (1988), though not all studies reach this conclusion. The diversity of the methodologies adopted amongst existing studies may explain the lack of consensus within the existing literature. In addition, multiple ability grouping practices have been evaluated though the strategies may not have homogenous effects on pupil attainment. It is also possible that the impact of ability grouping varies with the level of education, possibly with primary and secondary school setting having differential influences on attainment.

There are unities between the literature on ability grouping and the literature on peer effects since the ability grouping strategy adopted will determine the class peers of pupils. When grouping by ability, low attaining students are segregated from higher attaining students. The impact of ability grouping may therefore encompass the peer effect in addition to a teaching effect which impacts students through a change in the teaching style, pace, level or resources provided (Collins and Gan, 2013).

The peer effects literature generally suggests that the presence of higher ability pupils within the classroom environment benefits students who are less able (Kiss, 2013; Robertson and Symons, 2003; Bradley and Taylor, 2008) whereas in some cases, lower ability pupils are found to cause negative peer effects within the classroom (Lavy et al, 2011, Lavy et al., 2012). The peer effects literature therefore suggests that mixed ability teaching may benefit the lower ability pupils whilst higher ability pupils may benefit from ability grouping.

Gibbons and Telhaj (2012) examine the impact of peer effects upon student achievement when observing the transition from primary to secondary school, with an aim to identify whether pupils progress faster during their initial three years at secondary school should their schoolmates have performed well in KS2 national exams. Using data from the

National Pupil Database (NPD) and the Pupil Level Annual School Census (PLASC) the paper employs a value added and differencing based approach to estimate pupil progress between KS2 and KS3 national exams. By controlling for individual fixed effects alongside primary-by-year fixed effects and primary-by secondary fixed effects and trends, possible sorting and selection effects are removed, while controlling for unobservable characteristics and factors affecting students who make similar schooling choices. The study exploits the transition from primary to secondary school to utilize peer group reformation. The results of the analysis indicate that peer ability does affect achievement at age 14: a one standard deviation increase in the mean KS2 score of school peers leads to a 0.03 standard deviation increase in student achievement; this is however, a small effect suggesting that peer ability and group composition matters little. The paper notes that the results do not provide clear inferences for ability grouping.

With a greater focus on peer effects within the classroom, Atkinson et al. (2008) estimate the effect of more able peers on GCSE attainment in a UK setting, specifically in English and mathematics. The study adopts a methodology involving OLS, fixed effects and IV approaches using data from a unique dataset encompassing 9,428 pupils across two waves; the first, including pupils who sat the KS3 exams in 1997 and GCSE exams in 1999, the second including pupils who sat the KS3 exams in 2000 and GCSE exams in 2002. Due to the tier system of examinations[8] in English and mathematics at GCSE level, the authors argue that a setting policy is likely to be generated depending on how pupils are expected to perform and which tier of examination they will be entered into. Whilst the study considers a pooled estimation strategy, a within-tier estimation approach is additionally considered since it is argued that there is a more random allocation of students to sets within tier than within a whole school. In addition, in the pooled sample, students are not necessarily studying the same syllabus which may influence attainment. The study identifies that an increase in the KS3 class average increases attainment in all tiers, though the influence is reduced as moving from the higher tier to intermediate to foundation in maths. A small positive effect continues to be observed when observing the effect of class peer ability when allocation to classes is deemed as random. Overall, the study reports that for each subject, higher ability peers positively influence attainment; however, it is suggested that other factors may be influencing the results such as the

[8] In these subjects a number of 'tiers' are offered in the examination. The tiers will vary in difficulty. Usually on a higher tier exam pupils may obtain a higher maximum grade than on a lower tier exam. Pupils will be entered into one tier of the exam.

difference in exams, due to the tiered set-up of examinations, alongside differences in aspirations due to the tier, set and associated difficulty with gaining a "pass" grade.

Studies have alternatively analysed the impact of low ability peers; Lavy et al. (2012) utilizes English census data from 2003/4-2006/7 to observe the effect of peer quality upon KS3 test scores in the secondary school setting. The study measures peer quality using KS2 test scores alongside cohort-specific high and low-achieving proxies attained by identification of pupils in the bottom 5% of the national distribution of achievement. As in Gibbons and Telhaj (2012), peer group reformation is utilized through the exploitation of the transition from primary to secondary school. To overcome the potential bias due to selection and pupils sorting, the paper adopts a fixed-effects methodology while additionally exploiting variation in attainment across the core subjects: English, mathematics and science at KS3. The methodology involves observing whether the subject specific attainment variation between KS2 and KS3 for individuals is systematically associated with the subject specific variation in peers' ability. Defining students within the bottom 5% of the ability distribution as bad peers, the study finds that a reduction of bad peers from 20% to 0 results in an increase in KS3 test scores by 0.17 of the within-pupil standard deviation within the test score distribution. Overall, the study identifies an insignificant impact of average peer quality and 'good' peers upon pupil performance in the heterogeneous sample; however, when estimating the effect by gender, the presence of high ability 'good' peers is found to benefit girls; this effect is greater for girls at the bottom of the ability distribution. In contrast, an increase in the proportion of 'good' peers is found to have a negative impact upon the performance of boys, though this is an insignificant effect. Though the paper does present evidence that low ability peers hinder the attainment of schoolmates, including higher ability students, the impact of higher ability 'good' peers is less clear.

Though a share of the peer effects literature suggests that there is a benefit to mixed ability teaching, since higher ability peers are often found to have a positive effect, the literature does not explicitly compare mixed ability teaching to setting or other forms of ability grouping. Alongside peer effects, setting by ability is likely to involve other factors that influence outcomes and should therefore be considered when examining the impact of setting.

3.3.4 Class setting and non-academic outcomes

A small number of papers have additionally considered the impact of class setting and other forms of ability grouping upon a number of non-academic attainment outcomes, for example self-concepts and self-esteem (Abadzi, 1985; Ireson and Hallam, 2009), pupil attitudes (Boaler, 1997; Ireson and hallam, 2001; Suknandan and Lee, 1998) and grade anxiety (Wang, 2014). While a number of studies have investigated how pupil behaviour may determine set placement, few have investigated the causal impact of class setting upon behavioural outcomes, both as a net effect and as a differential impact dependent on the level of the set. The literature that estimates that impact of ability grouping on non-attainment outcomes is rather limited especially within the field of economics; additionally, there are few studies that adopt econometric techniques or complex quantitative analysis to overcome the surrounding problems thus making this section of relevant literature fairly limited.

Self-concepts are defined as the self-constructed beliefs that a person holds about himself or herself (Shavelson and Bolus, 1981). Ireson and Hallam (2009) explore the relationship between setting and grouping strategies within UK secondary schools and self-concepts, using data on a stratified sample of 23 secondary schools in the UK; the data provides information on the general self-concepts, academic self- concepts and student achievement for 1,600 14-15 year old pupils. Self-concepts and general concepts were taken from the Self-Description Questionnaire II alongside English, and mathematics self-concepts whilst an additional scale was created for science self-concepts. Adopting multilevel modelling, the study identifies that academic self-concepts, though not general self-concepts, are influenced by the extent of ability grouping within the school. Higher ability groups are found to have greater self-concepts than students in low-ability groups; this is true for English, mathematics and science.

The previously discussed research by Ireson et al. (1999a) which observed the impact of setting within 45 UK secondary schools additionally examined the impact upon the self-concepts and self-esteem of students. Self-esteem is measured using the Marsh Self-Description Questionnaire (Marsh and O'neill, 1984) and Rosenburg self-esteem scale (Rosenburg, 1965). Continuing to adopt multilevel modelling, the paper controls for pupil intake, through Year 6 (end of primary school) test score, gender and social disadvantage, measured by free school meal eligibility. The results from fixed effects and random effects approaches indicate that self-esteem and self-concepts are unaffected by setting in maths and science. Setting in English, on the other hand, is found to improve the self-concepts of low attaining pupils but lower the self-concepts of higher attaining pupils. These results mirror the findings of an early paper by Kulik and Kulik (1992) who undertake meta-analysis of 13 studies and also find no evidence of an influence of ability grouping upon self-concepts.

In addition to the peer effects literature that looks at the impact upon attainment, peer effects have also been related to behaviour. Carrell and Hoekstra (2010) attempt to estimate the peer effects associated with being in a class with children from troubled families, upon behaviour and reading and maths scores. Using US data, the paper exploits the variation in experienced domestic violence amongst children, as signalled by a court protection request, as exogenous variation in peer quality. Pupils from troubled families are found to exhibit more disruptive behaviour while pupils exposed to these children achieve lower academic outcomes and exhibit worse behaviour relative to their siblings who were not exposed to peers experiencing domestic violence. The study presents evidence that the behaviour of classroom peers is likely to influence both the pupils' own behaviour and academic outcomes. This may suggest that a pupil placed in a lower set, where peers are more likely to misbehave, is likely to exhibit bad behaviour themselves due to the influence of their peers. This may then have effects upon attainment and attitudes.

It is therefore likely that the setting of pupils by ability is likely to entail peer behavioural effects; since it is argued that the placement of pupils may be influenced by their behaviour (Dunne et al. 2007) there is likely to be a correlation between the overall class behavioural level and the set level; it may therefore be argued that the behaviour of lower set pupils is likely to be negatively influenced by the more probable deviant peers.

In summary, the evidence of the impact of ability grouping on non-attainment outcomes is rather inconclusive; of the evidence available, few studies adopt econometric techniques or employ complex quantitative methods as this chapter will do. Furthermore, the literature on the impact of ability grouping on student behaviour is limited; this is an additional gap which this chapter would like to fill. From the evidence, it seems that setting may be detrimental to the attainment of low ability students, particularly due to peer effects and possibly alternative factors such as possible lowered self-esteem. However, there is very little evidence or research on the impact of class setting upon behaviour thus there are few studies to which the results of this chapter may directly relate.

3.4 DATA

The Millennium Cohort Study (MCS) data is adopted within this chapter; the MCS is a national longitudinal birth cohort study which initially followed 19,000 children born in the UK between September 2000 and January 2002. To date, the MCS comprises of five waves, firstly when the children were aged 9 months in 2001/2 followed by a second wave in 2003 when children reached 3 years. The study provides data on the children biennially with the most recent wave corresponding to the interview years 2012/2013 when children were aged 11 years. By the fourth and fifth waves the study achieved samples of 14,043 and 13,469 children respectively.

This chapter will use data from the fourth wave, collected in 2008 when children were aged 7, alongside the fifth wave in order to analyse the impact of class setting upon behaviour. The analysis is based upon individuals from England and Wales only since the teacher survey, from which data will be utilised, was conducted in these countries alone.

The MCS provides a suitable dataset for this chapter since it provides a wealth of information on social, economic and health aspects of the children's lives. In addition, this dataset provides information on children who very recently attended primary school thus providing a rather current and up-to-date reflection of the policies adopted within schools.

A parent interview was issued in each wave to gain information on a number of topics such as the family context, the child's health, education and income, employment and parenting activities. Responses to the parent interview are provided by the main parent or carer; in around 98% [9]of cases this is the natural mother. Responses are additionally gained from the main parent's partner where applicable; in most cases this is the father of the child but may also be a different partner. From the second wave, at age 3, cognitive assessments and physical measurements were carried out on the child directly. From the fourth wave when children were aged 7, self-completion questionnaires were issued to the child respondents, covering topics such as family, friends, school and feelings.

[9] 97.4% in 2008 & 97.5% in 2012

In addition, the fourth and fifth waves of the MCS provide the responses to postal self-completion questionnaires completed by the child's teacher. These surveys cover questions about the child's abilities, behaviour, profile, parents, groupings, the characteristics of teacher and class and the move to secondary school. The teacher survey additionally provides information on the class set of the child through the inclusion of separate questions asking whether the child is set for mathematics, English / literature and science. The level of the set is also provided by the teacher who indicates whether the child is in the highest, middle or lowest set for each of these subjects. This chapter will focus specifically on class set in mathematics since this is the subject for which ability setting is more prominent amongst the MCS respondents and in primary schools more generally around the UK (Hallam et al., 2003); this is illustrated in Table 3.1 which shows that within the MCS, the proportion of children set for maths is much higher than in English in both years observed. It should also be noted that within the sample adopted, children are taught maths for an average of 5 hours a week when aged 7 and 5.3 hours per week when aged 11; this therefore accounts for approximately one day of teaching time per week. Children are therefore taught within maths sets for a considerable proportion of their school week[10].

Within the teacher and parent questionnaires, a definition of class setting is provided[11] thus reducing the potential problems in varying definitions of class setting across schools and teachers. The definition of streaming is additionally provided[12], again potentially reducing the issue of setting being reported when a different system of ability grouping is in place.

Also of particular interest to this chapter is the Strengths and Difficulties Questionnaire (SDQ) which is used to assess behavioural and emotional problems in children within the MCS. The SDQ is a behavioural screening questionnaire used broadly by psychologists, clinicians, educationalists and researchers[13]. The questionnaire is suitable for children aged between 3 and 16 years and may be completed by parents or teachers (SDQ info,

[10] The impact of setting in English was also considered and investigated but the results provided few interesting results, possibly due to the lesser time spent in English sets, relative to maths, and the smaller sample of pupils set for English.
[11] "Some schools group children from different classes by ability for certain subjects only and they may be taught in different ability groups for different subjects. We refer to this as setting." NatCen (2008)
[12] "Some schools group children in the same year by general ability and they are taught in these groups for most or all lessons. We refer to this as streaming."…"Other schools do not group children by ability between classes. Sometimes this may be because there are not multiple classes in the year" NatCen (2008)
[13] SDQ info (2016) reports that over 3,900 SDQ publications exist from across developmental, genetic, social, clinical and educational studies.

2014a). Within the MCS, teachers and parents alike were independently asked 25 questions relating to aspects of the child's behaviour (Goodman, 1997); these questions are given in Table 3.2. The responses to the 25 questions provide information on 25 attributes and the extent of these possible problems. The coding of the responses allows for the respondent to indicate whether the child shows signs of each behavioural or emotional issue by providing an answer from the options: 'not true', 'somewhat true' or 'certainly true'. Each SDQ attribute is recoded so that when it is reported 'certainly true' that the child exhibits negative attributes or behaviours, a higher value (equal to two) is given than when it is reported 'somewhat true' that the child exhibits a behaviour (score equal to one). A value of zero is given when the child does not exhibit a problem. Thus higher scores signal a greater problem. Attributes deemed positive are alternatively recoded so that children who do not exhibit a positive behaviour are given a higher score (equal to two) rather than a score of zero as with the negative attributes. Positive attributes that are 'somewhat' exhibited are given a lower score (equal to one) whilst those children who do are reported to 'certainly' indicate a positive attribute are coded zero since this does not signal a problem.

The SDQ responses provide information on five behavioural problem categories: emotional problems, conduct problems, hyperactivity/ inattention problems, peer relationship problems and prosocial behaviour (Gallop et al., 2013), with five questions for each category. Observing the responses to the five questions within each scale therefore indicates the extent of the problem measured overall by the scale, with a possible minimum score of zero and a maximum score of 10 since each individual question may have a value equal to a maximum of two. Whilst the SDQ may be analysed through the five individual categories, the SDQ problems may also be measured by a total difficulties score which sums the problems within the first four categories (emotional, conduct, hyperactivity and peer problems). The total difficulties score may therefore equal a maximum of 40 and provides an overall indication of the behavioural problems of children from all possible aspects. The total difficulties score is treated as a continuous variable (Coren et al. 2010).

Alternatively, the behaviour reported by the SDQ may be defined as internalising or externalising behaviour; emotional and peer problems may be summed to gain a measure of internalising behaviour whilst the problems within the categories of conduct and hyperactivity problems may be summed in order to gain a measure of externalizing behaviour. Each of these measures therefore has a minimum value of 0 and a maximum of 20; as with the total difficulties score, these outcome measures are treated as continuous. These measures provide two dimensions of how problems may be projected by an individual; often, externalising behaviours are targeted towards others or are an outward expression or manifestation. Internalising behaviour, on the other hand, describes more inward behaviours and expressions. Research has suggested that girls are more likely to exhibit internalising behaviours whilst boys are more likely to display externalising behaviours (McNeish and Scott, 2014; Leadbeateret al. 1999).

The measure of interest throughout this chapter will be the total difficulties score provided by both the teacher and the main parent (usually the mother); in an attempt to identify the source of the behavioural problems, the chapter will also explore internalising and externalising behaviours as outcome behavioural measures. In observing the dimensions of behavioural problems, the chapter may also be able to shed light on the differences in behaviour between girls and boys as a result of setting, in terms of the types of behavioural problems that they exhibit. All outcome measures will be observed for both teacher and parent responses to allow for comparison between the behaviour in the home and school environments. Observing both responses may also allow for any spillover effects of setting to be identified outside of the child's classroom or school environment. In total, there will therefore be six outcome measures explored throughout this chapter.

The fourth and fifth waves are specifically used within this chapter as these are the only two waves of the MCS which each provide information on the class set of the child. In addition, these two waves provide the SDQ responses from both the parent and the teacher allowing for the home and school behaviour to be analysed and compared. With all variables of interest provided within the two waves, the data suitably allows for the FE methodology to be employed.

When analysing the impact of the set level, only wave 5 is utilised since children are in the final years of primary schooling in which setting is more prevalent in UK primary schools (Hallam et al. 2003).This is indicated in the MCS data as presented in Table 3.1 where the prevalence of setting is much greater when pupils are aged 11 relative to when aged 7; wave 5 therefore provides a greater sample of set children. Observing set children only when aged 11 reduces the possibility that the policy of setting adopted by a school reflects other observable or unobservable characteristics which may then lead to biased estimates; instead schools that do set children are likely to do so in accordance with the general school policy trend of setting in later school years in the UK.

In wave five, only the teachers of respondents living in England and Wales were surveyed (Gallop et al. 2013). Since the methodology requires data from the teacher questionnaires from both waves 4 and 5, the chapter will be based on the analysis of individuals from England and Wales alone. Dropping respondents from Scotland and Ireland from the panel data formed of waves 4 and 5, leads to a reduction in sample size by 4,514 observations, from 18,682 to 14,476. This initial sample in both waves is based on all children who remain within the same school; the 2,539 children who do not remain in the same school as the previous wave are firstly dropped before the analysis, providing the initial sample sizes for each of the samples given in Table 3.3. Children are only observed if they remain within the same school since a school move is highly likely to involve a number of differences in school policies aside from setting, whilst moving school is also likely to influence upon aspects of the child's behaviour.

Individuals are additionally dropped from the sample when only one of either the teacher or parent reports the behavioural difficulties of the child; these individuals are dropped in order to maintain consistency in the sample and to achieve an accurate comparison of the teacher and parent results by observing the same individuals throughout.

Individuals are also dropped from the sample when the reported proportion of English as additional Language (EAL) children within the class is inaccurate; for some children within the sample the number of children in the class with EAL is reported to be greater than the actual class size. Though this does not apply to an excessive number of individuals (only 241), these are removed from the sample since this variable is used to instrument set placement within the analysis; accuracy is therefore important. The final sample size is provided in Table 3.3.

3.4.1 Weighting

Sampling weights are available and are applied due to the sampling design of the MCS. The sampling process of the MCS involved oversampling within smaller countries: Northern Ireland, Scotland and Wales. In addition oversampling was also carried out in areas of high child poverty and, within England, in areas with high populations of ethnic minorities. This method of oversampling allows for a suitable sample size to be obtained for those sub-samples that may otherwise fail to achieve a good response rate and therefore be underrepresented within the sample. Weighting adjustment is applied since a skew is likely to be introduced by the oversampling thus controlling for the sample weights provided in the MCS restores the original proportions of responses by sub-groups and populations within the sample. The weights applied throughout the analysis take account of the differential sampling alongside attrition and non-response (Centre for Longitudinal Studies, 2011). The UK country specific weight is used for wave 5 since the teacher questionnaire, and hence class set variable and SDQ scores, are only available for individual respondents within England and Wales when respondents are aged 11. The sample used within this chapter therefore is restricted to families and individuals who participated in the fifth wave of the MCS.

3.5 METHODOLOGY

This chapter is concerned with answering two main questions: Does setting influence behaviour and does the level of the maths set influence behaviour? To answer each of these questions two separate methodologies are adopted with alternative approaches being taken. In evaluating the impact of setting upon behaviour an Ordinary Least Squares (OLS) methodology is initially adopted with a fixed effects (FE) estimation approach later being employed to assist in overcoming the problem of unobserved heterogeneity. Waves 4 and 5 are utilised from the MCS when children are aged 7 and 11 respectively. To address if placement in the lowest maths set influences behaviour an alternative strategy is employed; alongside an initial OLS estimation approach, an instrumental variables (IV) technique is utilised in order to overcome a possible endogeneity issue; for this part of the chapter, children are analysed at age 11 only, as explained in the previous data section.

Despite the variation in methodologies throughout the chapter which could be separated into two parts, there are many similarities in the approaches taken in order to maintain consistency and to develop the analysis from the first question to the next. One way in which this is achieved is by maintaining the same outcomes measures of interest throughout the chapter; in both sections, the impact of setting is examined when observing behaviour measures by the total difficulties score alongside internalising and externalising behaviour scores. Relatedly, the outcomes measured by both the teacher and parent will be observed throughout; this approach allows for comparison in behaviour at home and at school which may possibly vary (Mitchell and Shepard, 2011; Lewis et al 2014). Furthermore, the analysis in both sections will firstly be undertaken when observing a pooled sample of individuals before splitting the sample by gender; evidence suggests that girls are more likely to exhibit internalising behaviours whilst boys are more likely to exhibit externalising behaviours (McNeish and Scott, 2014; Leadbeater et al. 1999). This chapter seeks to identify whether setting has a differential effect on males and females overall but also looks at the impact of setting upon the types of behaviour exhibited according to gender.

3.5.1 Does setting influence behaviour?

The chapter will firstly examine the impact of class setting in mathematics upon the behaviour of primary school children, as assessed by the SDQ measures. With data being available on the setting of respondents and SDQ scores in the most recent fourth and fifth waves, the impact of setting may be identified when children are aged 7 and 11.

When attempting to identify whether being set influences behaviour, an OLS approach will initially be taken in addition to a FE methodology. The OLS approach involves estimating the impact of setting when controlling for a number of school, individual and family characteristics. There is potentially a methodological issue in adopting this simplistic OLS model; since the outcome of interest is a measure of behaviour, it is likely that there are unobserved characteristics which are related to or determine an individual's behaviour, these may be intrinsic characteristics or personality traits that are likely to vary between each individual. Unobserved individual effects may therefore be correlated with the regressors. It is possible that different individuals' behaviour responds to setting in a different manner. Since these unobserved characteristics cannot be controlled for, the issue of omitted variable bias arises in adopting OLS due to unobserved heterogeneity; OLS estimates are therefore biased and inconsistent.

FE estimation provides a resolution to this problem. The fourth and fifth waves of the MCS both provide information on the setting of the child and other characteristics of interest, enabling the use of a FE methodology. The FE approach allows for the impact of a change in setting upon behaviour to be observed when controlling for school-level, individual and family characteristics; control variables are presented in Table 3.4, though the choice of controls will be discussed later within the methodology section 5.3

Consider the model where there are two years of data; t=1 and t=2.

(Eq. 3.1)

$$Y_{it} = \beta_1 X_{it} + a_i + \varepsilon_{it}\text{ , t = 1,2,T}$$

The term a_i captures the unobserved time invariant factors that affect the outcome of interest Y_{it} . ε_{it} indicates the error term. The subscript *i* indicates variation over individuals whilst the subscript *t* indicates variation over time. Averaging this equation over time for each individual *i* gives:

(Eq. 3.2)

$$\bar{Y}_i = \beta_1 \bar{X}_i + a_i + \bar{\varepsilon}_i$$

$\bar{Y}_i$ gives the average outcome for each individual i. Subtracting the second equation from the first, i.e. subtracting the averages from the original equation gives:

(Eq. 3.3)

$$Y_{it} - \bar{Y}_i = \beta_1 (X_{it} - \bar{X}_i) + \varepsilon_{it} - \bar{\varepsilon}_i\text{ , t = 1, 2, T}$$

This is equivalent to:

(Eq. 3.4)

$$\ddot{Y}_{it} = \beta_1 \ddot{X}_{it} + \ddot{\varepsilon}_{it}, \qquad \text{t} = 1{,}2{,}\text{T}$$

$\ddot{Y}_{it}, \ddot{X}_{it}$and $\ddot{\varepsilon}_{it}$, are the time demeaned data on Y, X and ε respectively. The FE model may be extended to include more explanatory variables and additional time periods. In all time periods ε_{it} is uncorrelated with explanatory variables thus the FE estimator remains unbiased.

FE modelling is utilised in order to overcome the issues associated with OLS. Unlike OLS, FE models essentially estimate an intercept dummy for each individual by OLS in order to model individual effects, such as the intrinsic characteristics or personality traits. The intercept therefore captures all differences among individuals. Since the FE model is concerned with estimating the effect of a change over time within each individual, time invariant explanatory variables are excluded from the model, thus characteristics such as race or gender are unobserved (Hill et al. 2008).

Applying the fixed effect estimation to the model adopted within this chapter:

(Eq. 3.5)

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Set_{it} + a_i + \varepsilon_{it}$$

Y_{it} denotes individual i's SDQ score at wave t, measured by either the total difficulties score, the internalising behaviour score or the externalising behaviour score. β_0 denotes the intercept. set_i is a dummy equalling one when the individual is set and zero otherwise. X_{it} denotes a vector of school, teacher, individual and family characteristics of individual i at time t. a_i is a set of fixed parameters representing all stable individual characteristics of individuals. Finally, ε_{it} indicates an individual and time specific error term.

It may be seen in Table 3.1 that there is a variation in the number of children set for Maths between the fourth and fifth waves of the MCS; thus between the ages of 7 and 11 there is an increase in the prevalence of setting among the respondents, since setting is likely to be implemented in older year groups in primary school (Hallam et al., 2003). It is also evident that variation also occurs in the behaviour of respondents over time given the SDQ measures as will later be discussed. The FE approach will therefore be utilized in order to identify how a change in being set for mathematics between two periods, when children are aged 7 and 11, influences a change in behaviour amongst children at primary school between the same two periods, when controlling for all other determinants of behaviour. A number of behavioural outcomes will be observed including the total difficulties score and internalising and externalising behaviour scores each provided independently by both the teacher and parent. In addition, the analysis will estimate the impact of setting initially on a pooled sample before estimating the effect separately for girls and boys to identify whether there is a gender differential in the impact of setting.

OLS estimates will be presented for comparative purposes throughout; the OLS controls are simply the same as those in the fixed effects model but in addition, the time invariant characteristics are controlled for in the former.

3.5.2 Does the level of the maths set influence behaviour?

This chapter is also concerned with identifying whether the level of the set that the child is placed into for maths, influences behaviour. Specifically, the analysis will investigate the impact upon behaviour of being placed in the lowest set for maths relative to similar non-set pupils but also relative to pupils in mid and high ability sets. The impact of setting upon the lowest set children specifically is analysed since the literature suggests that whilst the higher ability children may gain in terms of academic outcomes and self-concepts from setting, it is the lower ability children who are disadvantaged by setting. For this reason, lower ability children are often the centre of the setting debate.

The fifth wave of the MCS is utilized within this part of the analysis since setting is more prevalent in the older year groups within primary schools, around when pupils are in years 5 and 6 and thus aged around 9-11 years old. A higher proportion of children are therefore set in the fifth wave, when children are in the final primary years, than in the fourth wave as demonstrated in Table 3.1. Wave 5 data therefore provides a greater sample of set children.

Initially, an OLS model is estimated, controlling for time variant and invariant characteristics as presented in Table 3.4. The variable of interest is the low set binary variable which indicates whether the child is placed in the lowest maths set. The OLS model may suffer from endogeneity due to reverse causality; while behaviour is possibly influenced by set placement, the child's behaviour may influence the level of the set in which they are placed. This is since children with worse behaviour or more behavioural issues are more likely to be placed in lower ability sets (Dunne et al. 2007; Boaler, 1997). OLS will produce biased and inconsistent estimates of the impact of setting due to the violation of the OLS assumptions.

In a simplified model:

(Eq.3.6)

$$Y_i = \beta_0 + \beta_1 D_1 + \beta_k X_k + \varepsilon$$

'D_1' is endogenous if:

(Eq. 3.7)

$$E[Cov(D_1, \varepsilon \,|X)] = 0$$

In order to overcome this issue of endogeneity, an instrumental variables approach may be taken. There are two major assumptions of this approach; an instrument (z) should be correlated with the endogenous variable (D_1) but should be unrelated to the outcome variable:

(Eq. 3.8)

1.z should be uncorrelated with ε: $Cov(z, \varepsilon) = 0$

(Eq.3.9)

2. z should be correlated with D: $Cov(z, D_1) \neq 0$

Thus, the instrument z should affect the endogenous variable D but not the outcome variable Y directly once controlling for all X_k ; the outcome should only be affected by the instrument through the effect of the endogenous variable. It should be noted that the first assumption above is not testable unlike the second; economic theory is relied upon in order to establish the first assumption whilst the second may be tested by regressing D on z.

In an ideal situation, the impact of setting may be observed when an exogenous shock influences the set placement of primary school children; though this is not currently possible, by using instruments that exogenously change the likelihood of lowest set placement, this effect can be somewhat imitated. Applying the assumptions of the instrumental variable to this chapter, an instrument must be selected that is correlated with being placed in a low maths set at age 11 but should not be related with child behaviour as measured by the SDQ at the same age. The instruments adopted within this study are: the proportion of children who have English as an additional language (EAL) in the class, and the number of maths sets in the school year of the child.

The initial instrument, the proportion of children with EAL is calculated using the number of children in the class with EAL and the class size variable. Whilst setting for maths is a year group practise, the number of children with EAL in the general teaching classroom within a given year group should be considered fairly random, or of a similar size to the number of EAL children within a different class in the same year group. This is a suitable instrument since children with an alternative native language are understandably more likely to struggle within the classroom when being taught in English; Sammons et al. (2007) identified that EAL is associated with more cognitive development problems for primary aged children. Additionally, it has been noted that children with EAL are more likely to be overrepresented within lower sets whilst many primary schools adopted the practise of firstly placing EAL children in lower sets before language is improved (Dunne et al., 2007). A higher proportion of EAL children within the general class is therefore likely to influence the probability of the set placement of other children within the year group. It is likely that a higher proportion of EAL children increases the likelihood that a non-EAL child in the class is placed in a mid or high ability maths set. The proportion of EAL students within the respondent's class is unlikely to influence the behaviour of a pupil; the behaviour of pupils should not be affected by the proportion of children in the class whose native language is not English. Evidence suggests that EAL pupils do exhibit behavioural problems at age 3 and 5 though such problems are no longer apparent at age 10 (Sammons et al., 2007). It may therefore be suggested that since the behaviour of the EAL children specifically is not influenced by EAL status at age 10, the behaviour of peers should not be affected.

The second instrument, the number of maths sets within the respondent's school year, is also likely to be correlated with the likelihood of the pupil being placed within the lowest set; a higher number of sets within the school year makes placement within the lowest set for maths more unlikely since there are more alternative sets in which the child may be placed. The number of sets that the school allocates to each school year for maths is unlikely to influence the behaviour of the child. The number of sets may be related to school size which in turn is likely to be related with class size which evidence suggests may influence some child outcomes; however, class size is controlled for within the model. It seems reasonable to assume that the size of the school outside of the classroom has little influence on the SDQ scores of children, especially when considering the lack of evidence of school size effects.

The procedure for selecting an instrument is not straightforward since in small samples estimation by IV may produce biased estimates; there is the additional problem of weak instruments even when benefitting from a large sample. Furthermore, within this chapter, a number of outcome measures are adopted due to different dimensions of behaviour being observed. Moreover, three samples of data are employed when the pooled sample is split according to gender. These factors make the selection of an instrument slightly more complicated.

A rigorous instrument testing procedure was therefore undertaken to ensure the validity, relevance and strength of the instruments; the results of which are presented in Table 3.5. One initial step was to test the assumption that the instruments are correlated with the endogenous variable; this is known as the test of relevance. From the table it may be seen that in all models and samples the instruments are jointly significant. The instruments were additionally tested for individual significance; with the p-values of the proportion of EAL variable and the number of maths sets being equal to 0.01 and 0.00 respectively, both instruments are able to explain the placement within a low maths set.

Whilst the instruments must be correlated with the endogenous variable, they must also be uncorrelated with the structural error term. Since two instruments are used to instrument one endogenous variable, thus there are more instruments than endogenous variables, the instruments may be tested to see whether they are uncorrelated with the error term. Only when an equation is overidentified can the excluded instruments be tested for independence from the error term. This overidentification test signals the validity of instruments. Since all tests are insignificant even at the 10% level the null hypothesis of instrument validity can be accepted thus the model seems to be correctly specified.

One further test undertaken is a test for endogeneity of the variable that is instrumented; in this case, it was thought that placement in a low maths set or simply the set level may be endogenous due to set placement possibly being determined by behaviour. OLS estimation would be more efficient if the endogenous regressor is actually exogenous whilst using IV would sacrifice efficiency. The results given in column 4 within Table 3.5 suggest that though there are instances where endogeneity is identified, in some cases, the null hypothesis of exogeneity is not rejected thus the set level doesn’t seem to be endogenous in all models and samples. However, there is no clear pattern to when the set level is not deemed endogenous; there are instances in most samples where the set is found to be endogenous; for example, in the parent reported total difficulties model for the sample of boys, endogeneity is not identified though for the internalising behaviours reported by the parent, the null hypothesis of no endogeneity (exogeneity) is rejected. Since internalising behaviours contribute to the overall total difficulties score it seems that there may be a level of endogeneity relating to certain behaviours even though the tests may not identify this. In each sample, at least one behavioural measure is found to be endogenous whilst all the behavioural measures should be highly correlated since the internalising and externalising behaviours sum to give the total difficulties score. It seems plausible and possibly a cautious approach to continue with the IV methodology in all cases but also adopt an OLS approach for comparison and robustness.

A further test of validity is the F-statistic; Stock et al. (2002) suggest that an F statistic in the first stage regression that exceeds 10 may be deemed reliable when one endogenous regressor exists. From Table 3.5 column 5, it may be seen that for the pooled sample and for the sample of girls alone, the F-statistic is continuously over 10 for each of the outcome measures. For these samples, the instruments are seen as reliable and valid. However, for the boys sample the F-statistic narrowly fails to meet this criteria in two of the parent reported behaviour models where the F-statistic equals 9.9. It should be noted that the F-statistic to test for joint significance of the coefficients on additional instruments is always found to be significant thus the instruments have significant explanatory power for 'low set' once controlling for other exogenous variables. It may be argued that although the instruments perform well for the pooled and girls sample, the instruments are slightly weaker for the boys sample. This will be considered when evaluating the results.

The final statistic to note from within Table 3.5 is the partial R^2 which indicates the correlation between the instruments and the endogenous variable having partialled out the impact of other covariates. For the pooled and girls sample the partial R^2 is consistently above the boys sample. Again, this may suggest that the instruments are weaker for the boys sample.

One additional check that is made is to test that the instruments are unrelated to the behavioural outcome measures since the instrument should only influence the outcome measure through the endogenous variable that it is instrumenting; it should not have a direct effect on the outcome. In all cases, these unpresented results indicate that each instrument insignificantly influences behaviour.

To summarise, the instruments seem to perform well under the testing procedure and indicate validity, relevance and in most cases do not show any signs of the weak instrument problem. For the boys sample, the instruments do appear to be weaker than in the pooled and girls sample so this will be taken into account when interpreting the findings of the analysis. The instruments do not seem to indicate complete weakness for the boys sample so it is still important to compare the IV results with the OLS results for this sample.

The aim of this part of the chapter is to identify the impact of being placed in a low maths set relative to not being placed in a low maths set; this alternative therefore includes being set in a high or mid set but also not being set for maths. In order to achieve this, a binary variable was entered into the model to equal one when the child is in the lowest maths set and zero otherwise. This simply indicates the impact of being in the lowest set; however, there is likely to be an additional impact of actually being set rather than not being set for maths. A binary variable was therefore entered into the model to equal one when the child is set and zero otherwise. The combination of these two binaries therefore allow for the effect of lowest set placement to be estimated in relation to other sets and not being set. Consider the simplified model:

(Eq. 3.10)

$$SDQ = \beta_0 + \beta_1 SET + \beta_2 LOWSET + \beta_K X_K + \varepsilon$$

Where SET indicated whether the individual is set or not, LOWSET indicates whether the individual is in the lowest maths set and X_K simply indicates a vector of characteristics. The coefficient on SET will indicate the impact of being set for maths though this will specifically relate to being set in the mid or high level set, relative to not being set. The addition of the SET and LOWSET coefficients will be referred to as 'lowest set' and will indicate the impact of being set but additionally being placed within the lowest set for maths, relative to not being set. Thus the combination of the two coefficients acts like an interaction term though modelling in this manner allows the 'pure' effect of being placed in the lowest maths set to be observed, over and above the effect of being setted. The standard errors for the addition of the two coefficients are correspondingly estimated.

3.5.3 Variable descriptions and characteristic controls

The vector of characteristics (X_{it}) contains a number of school, individual, teacher and family controls which have been included due to the likely relationship with child behaviour, based on the child behaviour literature. The controls within the models are constant throughout this chapter though due to the nature of the FE model, time invariant controls drop out. Thus, the IV and FE models will maintain the same time variant characteristics. Whereas the FE analysis will focus on two waves of data, wave 4 and 5 of the MCS, the IV analysis will be based on the fifth wave only when children are aged 11 and setting in Maths is much more prevalent among this age group. The controls discussed will therefore reflect age 7 and 11 characteristics independently for the FE models whilst the same controls will measure age 11 time variant characteristics for the IV model, with the time invariant characteristics being measured at various points in time as will be discussed. A brief description of all controls is provided within Table 3.4 for reference.

Special educational needs (SEN) are controlled for in the models since learning difficulties and needs may be directly reflected in a child's behaviour. Additionally, SEN children may have specific educational needs or difficulties that necessitate additional learning resources or assistance. These additional resources may change their learning experience; Blatchford et al. (2009) identify a relationship between the support a child receives and their achievement, this support may additionally influence upon the child's behaviour. Furthermore, the behaviour of parents may vary when their child has SEN, for example through assisting more with homework or making adjustments to the child's lifestyle; Peters et al. (2007) identifies that the parents of a child with SEN were all more probable to feel very involved in their child's education.

Teacher years and teacher tenure control for teacher experience which may influence the ability or experience in controlling behaviour or working with children with behavioural issues; evidence suggests that teachers with many years of experience have significantly less control over student behaviour (Ritter and Hancock, 2007). Tenure is likely to be correlated with experience though it may additionally reflect the teacher's knowledge, comfort and familiarisation with school specific policy, both in a general sense and relating to behaviour. The class size and mixed year group variables control for the number of peers alongside the presence of older, more mature peers. Whilst evidence suggests that in smaller classes children are more engaged in learning and exhibit less disruptive behaviour (Finn et al., 2003), a significant influence of mixed year group classes has been identified upon prosocial and aggressive forms of behaviour (McClellan and Kinsey, 1999). These variables may also pick up the effect of other school characteristics related to school size.

Out-of-school activities have been found to be related to SDQ behaviour in a paper by Chanfreau (2015) when similarly using MCS data; breakfast club and after school clubs were found to have opposing impacts upon child behaviour, with after school clubs having a positive impact upon the total difficulties score. Breakfast club and after school club attendance are therefore independently controlled for.

The remaining parental response controls have been entered into the model in accordance with the behavioural literature; income for example is likely to pick up a number of effects including deprivation whilst low income has been identified as a risk factor in influencing child antisocial behaviour (Scott et al. 2010). In relation to this, attendance at parents evening alongside parent interest, as measured by the teacher, attempt to control for factors associated with parenting style which Scott et al. (2010) also find is associated with child behaviours.

Maths ability enters the analysis as a time invariant control indicating the score achieved within the maths assessment in wave 4 of the MCS when children are aged 7. Maths ability controls for underlying, intrinsic mathematical ability which is assumed to be constant over time. The MCS progress in maths[14] total raw score is used to control for maths ability since it provides an exogenous measurement of maths ability. The MCS score is used instead of the KS1 SATS score since the results from national examinations are known by teachers and are therefore likely to be used when determining the child's set. Additionally, the KS1 scores may reflect the influence of school, testing and peers, possibly including the impact of being set, due to the preparation for the SATS being in school.

The maths ability control enters the OLS model to analyse the impact of being set but as underlying, intrinsic maths ability is assumed constant and time invariant, it does not enter the FE analysis. When attempting to answer whether the level of the maths set influences SDQ scores, the chapter takes OLS and IV approaches both including and excluding the maths ability score. In including ability in the model, differences in behaviour between set and not set cannot be due to ability; this would mean separating out the effect of setting from the effect of ability so that the impact of set is not reflecting the ability level. This may reduce the potential for a confounding variable or omitted variable bias to arise. However, the placement within a set by is likely to be determined by ability whilst ability is also likely to play a part within the level of the set impacting upon behaviour. Analysis is therefore undertaken both including and excluding the ability control.

The remaining time invariant controls are entered into the analysis to reflect the determinants of child behaviour in accordance with the relevant literature. These controls are measured in different years; white, date of birth, birth order and birth weight for example are taken from the first wave of the MCS when children were 9 months old. It is important to control for factors such as ethnicity since evidence suggests that the teacher reported incidence of abnormal and borderline behavioural problems amongst children

[14] The MCS progress in maths score provides a measurement of mathematical ability which indicates progress in relation to the National Curriculum in the UK. The test undertaken by the MCS respondents is a reduced version of the National Foundation for Educational Research standard Progress in Maths (PiM) test. The test is undertaken at age 7 in the fourth wave of the MCS and involves a series of 'paper and pencil' calculation exercises covering a number of mathematical topics (Connelly, 2013). The total raw score is used within this study as a control for ability, this simply represents the number of correct answers given on the test.

varies by ethnicity (Popli & Tsuchiya, 2014). Similarly, the season of birth is identified as being a determinant of behaviour, whilst it is argued that younger children may exhibit behavioural immaturity which may result in lower set placement due to perceived lower ability by teachers (Campbell, 2013). Birth order and birth weight are additionally controlled for; whilst evidence suggests that birth weight is correlated with the susceptibility to issues such as anxiety, depression and aggressive outbursts amongst school children (Bohnert & Breslau, 2008), the existing literature also identifies a significant influence of birth order upon child behaviour and the behavioural roles adopted (Sulloway, 1996).

Parental education is controlled for by tracking the highest qualification obtained from wave one to wave five to identify whether a parent has a degree; very few individuals obtained a degree between the observed years, thus parental education will drop out of the FE model, since being time invariant.

Attendance at a religious service is controlled for by observing whether the child attends a service less than once a year; this is likely to reflect a family's involvement in religion. This information is only available in wave 5 but is assumed to be fairly constant over time since a level of interest, commitment or underlying belief to a religion is likely to remain while the frequency of attendance may vary. Attendance is more likely to be determined by parents in the household who may be likely to pass on beliefs or values to children which may influence behaviour. In addition, a sense of identity or belonging may be gained from being part of a religion; this may influence the child's values, outlooks and attitudes, in turn potentially influencing behaviours measured within the SDQ, as identified by Petts (2009). It could be assumed that such effects do not disintegrate with service attendance.

3.6 RESULTS

3.6.1 Descriptive statistics

Table 3.6 provides descriptive statistics relevant to the initial analysis within this chapter where the impact of being set in Maths upon behaviour is investigated and a FE methodology is employed. Variation in setting and the covariates over the observed time period is central to the FE model; table 3.6 indicates that between the ages of 7 and 11, respondents are likely to experience a change in setting, while there is considerable variation in many of the characteristic controls. This variation over time is indicated by the within standard deviation and is substantial for characteristics such as parent evening attendance, indicating that attendance changes for individuals over time. This is also true of other continuous measures, for example teacher years and class size; thus, children are likely to experience different levels of teacher experience and class sizes over time. Table 3.6 also provides other information of interest, such as the mean of the individual covariates, relevant for the FE analysis.

Similarly, Table 3.7 presents the summary statistics for the outcome variables: the total difficulties score, the internalising behaviour score and the externalising behaviours score, provided by both the teacher and the parent. Individuals experience a greater variation in the total difficulties score between age 7 and 11; this is not surprising since the internalising and externalising scores reflect a smaller number of components within the total difficulties score, as described in Table 3.2. The total difficulties score variation thus indicates the variation in both internalising and externalising behaviour over time. The statistics suggest that over the two periods observed, the teacher reported total difficulties score is likely to change to the greatest extent, relative to parent reported scores, possibly since the child's teacher may change over time, their parents do not. In addition, internalising behaviours seem to vary to a greater extent within individuals than externalising behaviours.

In relation to the variation in characteristics over time, Table 3.8 provides the statistics on individuals' experience of setting between waves 4 and 5. A large proportion of individuals within the sample experienced no setting for maths when aged 7, but were set when aged 11. This is the most common experience of setting in maths within the sample, followed by not being set for maths at all over the observed time frame. Data is therefore provided on 2,091 individuals who experience a change in setting over the time period of interest.

Figures 3.1 and 3.2 provide a visual representation of the distribution of the SDQ scores for both the teacher and parent responses in 2008 and 2012. It is clear that in all cases, the distribution of the total difficulties score is skewed to the left, thus higher scores approaching the maximum of 40 are less common; very few individuals are reported to have scores greater than thirty in each case whilst individuals are commonly reported to exhibit a score between zero and ten. One interesting feature is that the mean difficulties score reported by both the teacher and the parent increases over time i.e. the mean difficulties is higher when respondents are aged 11 (in 2012) relative to when aged 7 (in

2008); the mean teacher reported total difficulties score increases from 5.9 to 6.8 whilst the mean parent reported total difficulties score increases from 7 to 7.3 between 2008 and 2012. In addition, teachers seem less likely to report no behavioural problems when children are aged 11 (in 2012). Additionally of interest is the difference in the distribution and mean of teacher and parent reported scores; the mean score reported by the parent is higher than that reported by the teacher. This may possibly be due to parents being more likely or able to recognise problems within their own child whilst parents may also spend more time with the child than the teacher. Over the two periods, the increase in the mean behavioural problems reported by the teacher is much larger than the change in the mean of parent reported problems.

Figure 3.1 Total difficulties by year – Teacher reported

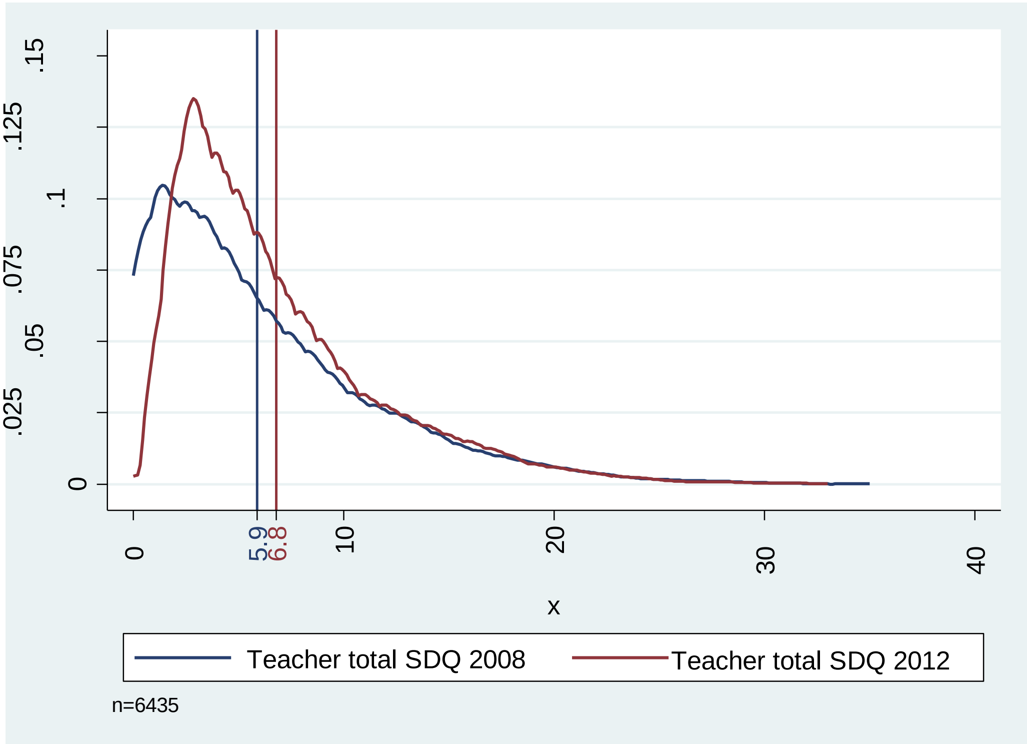


Figure 3.2 Total difficulties- Parent reported

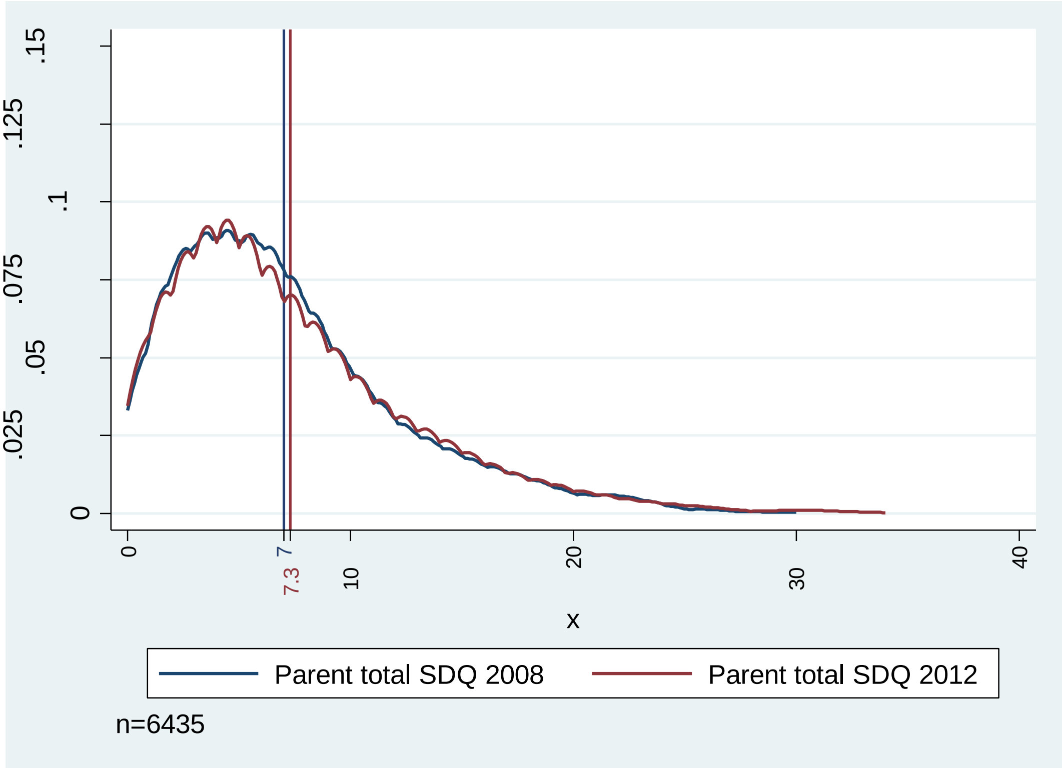


Table 3.9 provides the percentage and frequency of individuals within each set level for maths alongside those who are not set for maths; this is relevant for the analysis of the impact of being placed in the lowest maths set. Though there are fewer boys reported to be placed within the lowest maths set, there is not a large difference in the proportion of girls and boys placed within each set level. The table also indicates that over 39% of respondents are not set which is a larger proportion than expected given figures from Table 3.8; this is due to missing data for some individuals at age 7, who were not set at age 11. This is another reason for using 2012 data alone for the analysis of the level of set placement.

Descriptive statistics for the 2012 (age 11) specific covariates may be found in Table 3.10 and are relevant to the second part of the chapter. These descriptive statistics of the full sample in Table 3.10 may be compared to when the sample is restricted to the lowest set individuals, provided in Table 3.11, to highlight differences between all individuals and low set individuals. Some notable differences between the statistics include ability, which would be expected since ability is a large factor contributing to the setting of children in maths. Additionally, the mean values of SEN indicate that there is a higher proportion of SEN children within the lowest maths set sample, relative to the sample overall. There is also a notable difference in the parental degree variable as there is a lower incidence of degree educated parents within the lowest set sample. Parental education may be associated with a number of factors which potentially influence set placement through influencing the child's educational attainment; for example, help with school work, household resources and parental interest may vary between children with educated and uneducated parents. The tables also indicate that lowest set children are less likely to be the first born, white and born in the autumn or winter, all as expected. On average, the lowest set are also less likely to have married parents and live in a working household than the sample as a whole; parents are also less likely to be deemed 'interested' in the child's education as reported by the teacher.

Similarly allowing for comparison between the full sample and the lowest set sample, Table 3.12 provides the descriptive statistics for the outcome measures of interest: the total difficulties score alongside the internalising and externalising behaviour scores[15]. It is clear that those in the lowest maths sets exhibit worse behavioural problems than children in other sets or who are not set. This is true for boys and girls alike. It could be argued that this also presents evidence that individuals who exhibit worse behaviour are placed in lower sets though the raw statistics do not present a causal relationship. These data may highlight the issue surrounding the possible endogeneity problem in estimating the impact of set placement, which is addressed within this chapter. The table also shows that girls seem to exhibit fewer behavioural problems than boys, yet they are just as likely to be placed within the lowest maths set. Additionally, the mean number of parent reported problems is consistently higher than the mean number of problems reported by teachers for the pooled sample and the girls sample, as also indicated in Figures 3.1 and 3.2 and in Table 3.7. In contrast, for boys, the teacher reported problems are greater than parent reported behaviours.

3.6.2 Results: Does setting in Maths influence behaviour?

Table 3.13 provides a summary of the results from the examination of whether being set in maths influences upon child behaviour. OLS and FE results are presented for all models of both parent and teacher reported behaviours. OLS results indicate a predominantly positive influence of setting upon behavioural problems. The results do not reveal any significant role of setting in maths upon teacher or parent reported total behaviour scores. When alternatively looking at internalising and externalising dimensions of behaviour, for boys, the results reveal a weakly significant positive influence of setting upon teacher

[15] The distribution of the scores for the pooled sample may be observed in the previously discussed Figures 3.1 and 3. 2.

reported internalising behaviours whilst for the pooled and girls sample, the results indicate a positive and significant impact of setting in maths upon parent reported externalising behaviours, thus suggesting that setting may be detrimental to this dimension of behaviour since increasing the reported behaviour scores.

The FE results which attempt to correct for the unobserved heterogeneity which may bias the OLS results indicate that there is a positive and significant influence of setting when observing the pooled sample; ceteris paribus, being set in maths increases the number of teacher reported total difficulties score by 0.32 relative to when not set for maths. This is a significant effect but at the 10% level only. A similar finding is identified for boys but the effect is of a greater magnitude; the total difficulties score of boys who are set for maths is 0.69 higher than boys who are not set, ceteris paribus.

The results also identify that setting positively impacts the teacher reported internalising behaviour score; for boys, being set for maths increases the internalising behavioural score by 0.52 ceteris paribus; this is a highly significant effect indicating that setting in maths is detrimental to the behaviour of boys, specifically internalising behaviours which relate to emotional and peer related difficulties. A similar detrimental impact upon behaviour is identified in the pooled sample, though the result is significant at the 10% level only.

Interestingly, the results suggest that the behaviour of girls is improved by setting in maths; the parent reported internalising behaviour score of girls who are set is 0.30 lower than girls who are not set, ceteris paribus. Thus for girls, emotional and other forms of inward directed behaviours improve when set for maths.

Though the effects differ in sign, the results indicate that similar dimensions of behaviour of boys and girls are affected by setting. The existing literature often finds that boys are likely to exhibit externalising behaviours whilst girls are more likely to exhibit internalising behaviours (McNeish and Scott, 2014; Leadbeateret al. 1999); these results indicate that internalising behaviours are influenced by setting in maths for both boys and girls, though varying by who reports the behaviour.

Since the impact of setting includes children from all sets of maths, the effects are difficult to disentangle and provide rationale for the secondary analysis to be undertaken within this chapter to analyse the impact of the level of set placement. Based on conjecture, this effect could be explained by the difference in confidence between girls and boys; an

OECD (2014) report identified that there were large differences in the attitudes towards maths for girls and boys; unlike boys, girls were found to suffer greater anxiety in maths, have lower self-confidence and lower belief in their own ability. Emotional aspects of behaviour for girls may improve when the class caters for their ability level thus reducing the chance of withdrawal or knocked confidence when the class work is inaccessible or difficult to understand. Equally, higher set girls may gain confidence from their set placement. Boys on the other hand who may not struggle initially with 'maths anxiety' may be more influenced by the actual placement; for example, it could be that the confidence of lower set boys is knocked by being placed in lower sets whilst the more difficult classes for higher sets have a similar impact on emotional aspects of behaviour when work is more challenging.

The results vary between the OLS and FE models; for teacher reported behaviours in particular,the OLS point estimates are often smaller than the FE estimates. The OLS models generally suggest a positive influence of setting whereas the FE results identify both positive and negative impacts of setting dependent upon the sample, though results are predominantly insignificant.

The full FE results which may be found in the appendix (Tables A3.1-A3.2) additionally provide interesting results. SEN is found to have a persistent positive and significant influence upon most dimensions of behaviour both in the OLS and the FE models. Children with SEN exhibit more behavioural difficulties than non-SEN children whilst a change in SEN status over time also influences behaviour; this may be due to the additional resources and assistance provided to a child when receiving SEN status, which may alter their behaviour. The findings may mirror those of Fauth et al. (2014) who identify with MCS data, that over time, children with SEN encounter a greater increase in peer, hyperactivity and emotional problems, as identified by the SDQ, relative to non-SEN children; the behavioural trajectories of SEN children relative to non-SEN children are therefore likely to diverge over the primary school years.

Additionally, parental interest is found to be a rather consistent significant determinant of teacher reported behaviour in a number of the FE models; children with interested parents exhibit fewer behavioural difficulties which could possibly be due to more interested parents being more likely to identify and attempt to solve problems. Relatedly, a child with interested parents may be less likely to initially exhibit behavioural problems such as externalising behaviours which may often involve aggression or hyperactivity which may be an effort to gain attention.

3.6.3 Results: Does lowest set placement influence behaviour?

Table 3.14 provides the OLS results from the examination of the impact of lowest set placement upon behaviour. The coefficient on ‘low set’ indicates the impact of being in the lowest maths set, relative to individuals in other sets and also those who are not set. The coefficient on ‘set maths’ indicates the impact of being set for maths though this specifically relates to being set in the mid or high level set, relative to not being set. The addition of the coefficients on ‘set maths’ and ‘low set’ is given by ‘lowest set’ and indicates the impact of being set but additionally being placed within the lowest set for maths, relative to not being set.

The results indicate that being placed in the lowest set for maths is detrimental to all dimensions of behaviour for the pooled and girls sample. For boys, the effect is not consistently significant, though there is strong evidence for an impact of lowest set placement upon the teacher total and internalising score alongside the parent reported externalising scores.

The ‘set maths’ coefficient may be interpreted as the impact of being set when placed in alternative sets, i.e. being set in middle or high sets, relative to being taught in a mixed ability class (not set). Throughout the OLS results the impact of being placed in a mid or high set is identified as having a negative effect upon all dimensions of behaviour, though this effect is not consistently significant in all models. Thus, setting in maths seems to be beneficial for the behaviour of mid and high set children whilst for the lowest set children, setting in maths is damaging to behaviour. These initial OLS results therefore mirror the findings of a number of studies within the surrounding literature (Hallam and Parsons, 2014; Gamoran, 2002; Suknandan and Lee, 1998; Slavin, 1988; Ireson, 1999a).

Table 3.15 provides the subsequent IV analysis results. The pooled models indicate that there is an insignificant influence of being placed in the lowest set for maths upon all dimensions of behaviour. In the pooled sample, the behaviour of children in the lowest set is therefore insignificantly different from the behavioural scores of children that are not set and are therefore taught in mixed ability classes for maths. These results mirror the findings of Ireson and Hallam (2001) and Whitburn (2001) who identify little evidence of an effect of ability grouping upon achievement outcomes.

For girls, as in the pooled sample, the total behavioural score provided by both the teacher and parent is found to be insignificantly influenced by both setting and being placed in the lowest set for maths. However, when alternative dimensions of behaviour are investigated a beneficial impact of lowest set placement is identified; ceteris paribus, girls who are placed in the lowest set for maths are reported to have an internalising behavioural score that is 2.3 lower than girls who are not set for maths. This is a significant effect at the 5% level which remains significant at the 10% level when ability is included within the model though the magnitude of the effect falls slightly. The effect of being placed in a mid or high set upon behaviour is found to be insignificant, thus, rather than benefitting from setting, the behaviour of high and mid set girls is unaffected. The results provide some support for the FE results that, in some models, suggested that the setting was beneficial for girls' behaviour.

This change in internalising behaviour for girls supports a wealth of literature which suggests that girls are more likely to exhibit internalising behaviours (McNeish and Scott, 2014; Leadbeateret al. 1999). These internalising behaviour problems are often explained to be associated with social withdrawal, attention seeking, dependency and feelings of worthlessness (Guttmannova et al. 2008). By reducing these behavioural issues within girls, setting may help students to be more engaged within the class, reducing the likelihood of lower ability children becoming withdrawn from the class and 'acting up' (House of Commons, 2011) when the class is beyond their understanding. By providing teaching that targets the ability of the class, the material may be more accessible and comprehensible, allowing girls to gain confidence; as explained in the FE results discussion, this may be particularly important for girls in maths. Furthermore, the removal of the highly able children from the class may reduce the pressure and stress caused by comparing ability amongst classmates.

For boys, a largely insignificant impact of lowest set placement is identified though a weakly significant positive effect is identified upon parent reported internalising behaviour; relative to boys who are not set for maths, being placed in the lowest set leads to a 3.69 increase in the internalising behavioural score of boys, ceteris paribus. The magnitude of this effect increases slightly when controlling for ability. Being placed in middle or high sets relative to being taught in a mixed ability class reduces the number of behavioural problems reported by a parent by 0.9 problems, ceteris paribus. Thus, while the behaviour of boys is improved when placed in a high or mid set for maths, placement in the lowest set for maths is detrimental to behaviour, in accordance with the findings of the existing literature (Hallam and Parsons, 2014; Gamoran, 2002; Suknandan and Lee, 1998; Slavin, 1988; Ireson, 1999a), though the results provide rather weak evidence. As in the FE analysis, internalising behaviours are influenced by setting for boys whilst the literature suggests that externalising behaviours are more likely to be exhibited (McNeish and Scott, 2014; Leadbeateret al. 1999).

The results therefore provide some evidence that the behaviour of boys deteriorates when placed in lower sets. This may be due to peer effects within low sets as behaviour is likely to be correlated with the level of the set (Dunne et al. 2007). Pupils may then be influenced by the behaviour of their peers and subsequently also adopt bad behaviour, possibly also at home (Carrell and Hoekstra, 2010). Another explanation may be that, the self-concepts and self-perceptions of children are diminished when placed in lower sets; Kutnick et al. (2005) found that children may be demotivated by lower level setting whilst Campbell (2013) argued that children may behave in correspondence with the role assigned to them. The effect identified is reported by the parent but not teacher, thus it is possible that self-esteem and confidence are impacted which may be projected by behaviour at home.

The IV results vary to a large extent with the OLS results which identified a consistently positive impact of lowest maths set placement on behaviour; the IV results do not provide such strong evidence for a significant role of set placement in determining behaviour. Thus, when addressing the issue of endogeneity which may bias the OLS results, the findings across the OLS and IV models are inconsistent.

The full IV results[16] which are given in the appendix (Tables A3.3-A3.5) also indicate some other interesting determinants of child behaviour. As in the FE analysis the results from most models indicate a large and significant role of parent interest in determining child behaviour; for example, the teacher reported total difficulties score of children with parents who are deemed interested is 3.1 lower than children with 'uninterested parents', ceteris paribus. Similarly, SEN is a consistently positive and significant determinant of behaviour across all models, as also identified within the FE analysis. In the pooled sample, having SEN increases the number of total difficulties reported by the teacher by 4.19, ceteris paribus (Table A3.3). This is a large effect but is understandable since SEN may include children with difficulties that encompass behavioural issues. The analysis will later be undertaken when SEN children are excluded from the sample to identify whether the inclusion of these children in the sample is influencing these results; it is

[16] Full results are given for all pooled models and the total difficulties score for the gender results

important to identify whether the behaviour of children that is not explained by specific difficulties is impacted by setting.

The season of birth is additionally found to influence the behaviour of both boys and girls; in the pooled sample, being born in the autumn or winter leads to a 0.34-0.38 reduction in the total difficulties score, ceteris paribus. This is not a surprising finding given the existing literature; children who are born later in the year are often identified as performing worse and exhibiting worse behaviour due to their relative immaturity (Campbell, 2013).

Attendance at breakfast club is found to insignificantly impact upon all dimensions of boys' behaviour but positively impact upon the teacher reported behaviour of girls; attendance at breakfast club increases the number of teacher reported behavioural problems by around 0.47 for girls, ceteris paribus. Though these results oppose the results in the existing literature which suggest that such clubs may improve social skills, social competence and reduce the possibility of isolation which may of course benefit the behaviour of the child (Barker et al.,2003), these results may agree somewhat with the findings from Chanfreau (2015) who similarly identifies that out of school activities such as attendance at after school clubs may increase the total difficulties score of children as reported by parents within the MCS. The study identifies that the reasons for attendance may alter the impact upon SDQ outcomes since children may attend for child care reasons rather than actually wanting to attend which may influence their behaviour.

One other factor that influences one gender to a greater extent is class size; for girls class size has a weakly significant influence upon behaviour, in a limited number of models. For boys, class size is a stronger determinant of behaviour, influencing both teacher and parent total difficulty scores. An increase in the class size by one reduces the parent reported total difficulties score by 0.06-0.07, ceteris paribus. Thus the behaviour of boys is improved with a greater class size.

Living in a working household is identified as a negative determinant of behaviour, predominantly in the internalising behavioural models. The pooled model results show that children living in a working household have a teacher reported total difficulties score 0.7-0.8 lower than children in non-working households, ceteris paribus. It could be postulated that living in a household with a working parent may provide motivation to children to do well in school in order to similarly gain employment in the future. As significant effects are identified upon internalising behaviours the findings agree with the results of Hope et al. (2014) who identified that children were at greater risk of socio-emotional behavioural problems when not having an employed parent for one or more wave of the MCS, relative to having a continuously employed parent.

Children of degree educated parents are consistently found to exhibit fewer behavioural problems; for girls, fewer parent reported difficulties are identified whilst for boys, fewer teacher reported behavioural issues are reported. Parental education may be correlated with parenting styles which Scott et al. (2010) find is associated with child behaviours. Educated parents may provide more help and assistance with homework and tutoring which may be reflected in improved behaviour while educated parents may be more likely to attempt to tackle behavioural problems.

Having a regular bedtime is found to have a beneficial impact upon the behaviour of both boys and girls in a large number of models. The magnitude of this effect is greater for boys who, when having a regular bedtime, have a parent reported total difficulties score between 1.1 and 1.3 lower than boys with an irregular bedtime. This identified effect of a regular bedtime improving behaviour corresponds with the existing literature (Kelly et al., 2013).

Having married parents is a consistently significant and negative determinant of behavioural difficulties; in the pooled sample, the results indicate that the total difficulties score of children with married parents is 0.5-0.7 lower than children with a single parent, ceteris paribus. Platt (2014) similarly identify that this relationship between family structure and behaviour exists when adopting MCS data. Having a two parent household may reduce behavioural problems since care is shared between two parents whilst behavioural problems may be dealt with more easily with secondary support. Additionally, a family break-up may influence the behaviour of a child.

The IV results also indicate that an increase in the number of siblings in the household reduces the parent reported behavioural scores of girls. This could arguably be due to parents and children having less time to spend together when more children require care, therefore, parents may be less aware of behavioural problems exhibited by their children. The behaviour of boys, on the other hand, is insignificantly influenced by the number of siblings in the household. It could be argued that since the behaviours of girls are more likely to be internalising, inwards forms of behaviour that having less time to spend with the child makes these behavioural problems more difficult to identify whereas for boys who usually exhibit more externalising behaviours, these problems are easily identifiable even in large households with many children since externalising problems relate to issues surrounding hyperactivity, aggression and other forms of outward behaviours.

Similarly influencing the behaviour of one gender only is birth order which is found to influence the behaviour of boys but not girls; an increase in birth order by one, reduces the parent reported total difficulties score of boys by between 0.4 and 0.5, ceteris paribus. Being born later relative to other siblings leads to fewer behavioural problems. Parent reported behaviours only are influenced by birth order. This may be an unexpected finding since within the existing literature a later birth order has often been related to factors such as childhood and adolescent depression (Reinherz et al., 2003) though this may reflect family size.

One final notable finding is that family income influences the behaviour of girls to a greater extent than boys since income is a consistent negative determinant of parent reported behaviours for girls; family income is often associated with favourable parental practises but also household assets which seem to influence the girls to a greater extent than boys as identified by Deng et al. (2013).

3.7 ROBUSTNESS CHECKS

A number of checks are carried out to test the robustness of the main results. Firstly, maternal depression is included within the parent reported behaviour models; this check is carried out in both the OLS and FE analysis which looks at how setting influences behaviour, alongside the OLS and IV analysis which analyses the influence of lowest set placement.

As an additional test of robustness, the lagged reported behaviour score is included in the model for the level of set analysis (OLS/IV). The maternal depression and lagged behaviour controls are included as robustness checks since the inclusion of these controls leads to unequal sample sizes between the parent and teacher models which should be comparable.

An alternative robustness check involves excluding children who are deemed to have much greater behavioural issues are from the sample to identify whether children without specific behavioural difficulties are influenced to a different extent to the main results.

Finally, as an extension of the lowest set placement analysis and as a comparative exercise, the impact of being placed in the highest set for maths is analysed in the same manner to placement in the lowest maths set.

3.7.1Maternal depression

Existing research has identified that depression causes a negative perceptual bias which leads to depressed parents overstating the behavioural problems of their child, relative to teachers and self-reports (Mowbray et al. 2005; Biggs-Gowan et al. 1996; Leis et al., 2014). Within this literature, the accuracy of parent reported measures of child behaviour is therefore questioned since depressed parents may produce exaggerated indicators of the child's internalising, externalising and total behaviour problems. It is therefore important to identify whether the inclusion of parental depression within the models of this chapter generates results which vary from the main results, which do not initially account for this potential cause of over reporting.

Since within the MCS, parent responses are gained from the mother, maternal depression specifically is controlled for. The maternal depression variable relates to whether the mother is currently being treated for depression. Maternal depression is firstly entered into the parent reported behaviour models of the FE analysis to control for any changes in maternal depression over time which may then impact upon reported behaviour. Similarly, the control is entered into OLS and IV models in the second part of the chapter when examining the impact of the level of set placement. The inclusion of this control produces a difference in sample sizes between the teacher and parent models; as indicated in Tables 3.16 - 3.18, the FE and IV analysis achieves a sample size which is slightly smaller than in the main models given in Tables 3.13 – 3.15. Throughout the three models, maternal depression is identified as a positive and significant determinant of the parent reported behavioural score, suggesting that having a mother with depression leads to an increase in the number of behavioural problems reported. Similarly, a change in the 'depression status' of a child's mother over time has a similar influence upon reported behavioural problems, as reported by the FE models.

Table 3.16 gives the results of the re-estimation of the FE model including maternal depression. The results of this FE model indicate that being set for maths influences the parent reported behaviour of boys, specifically the total difficulties score; relative to not being set for maths, boys who are set have a 0.75 higher total difficulties score, ceteris paribus. Though this effect varies with the main results which indicated that setting did not influence parent reported behaviour for boys, the effect is significant at the 10% level only.

The inclusion of maternal depression in the model indicates that there is an insignificant impact of setting upon the parent reported internalising behavioural score for girls which additionally contrasts with the main results. However, the point estimate is almost identical across the two models; the effect of setting may be insignificant when maternal depression is controlled for due to the smaller sample size and consequently the higher standard error.

The results for the IV analysis is given in Table 3.17; the results indicate a positive and significant impact of being placed in the lowest set for boys; boys placed in the lowest set for maths have a parent reported internalising behavioural score of 3.26 greater than boys who are not set for maths, ceteris paribus. Being placed in the mid or highest set reduces the behavioural score for boys. These weakly significant results are consistent with the main model; the parent reported internalising behaviours of boys continue to be worsened by being placed in the lowest set whilst for girls, parent reported behaviours are insignificantly influenced by set placement.

For conditioning on depression to affect the setting coefficient, maternal depression would have to be correlated with the setting variable. For the IV analysis, it could be argued that this may be possible since maternal depression may hinder a child’s cognitive development (Canadian Paediatric Society, 2004) thus possibly influencing the level of set placement. For the FE analysis, however, there may be few arguments to support how maternal depression may influence whether the child is set or not for maths.

3.7.2 Lagged behaviour score

The set in which a child is placed is likely to be influenced by their behaviour prior to the setting process. In addition, behaviour in a previous period is likely to influence behaviour in the next period since the behaviour of the same child is likely to follow a course or pattern; it is expected that few children with low behavioural scores when aged 7 would have great behavioural issues when aged 11 without a shock to the child. It is therefore interesting to observe whether setting continues to influence behaviour once previous behaviour is controlled for.

The lagged behaviour score will be controlled for in the IV model by including the behaviour score in the previous wave of the MCS (wave 4), when respondents were aged 7.

Table 3.18 provides the IV results that indicate that when controlling for lagged behaviour scores, being placed in the lowest set no longer significantly influences internalising behaviours for girls or boys. However, externalising behaviours are impacted by being placed in the lowest set; for boys, being placed in the lowest set for maths reduces the teacher reported externalising behaviours score by 5.8, ceteris paribus. This weakly significant effect is only present when ability is controlled for in the model. Parent reported externalising scores are reduced by between 3.6 and 4.1, ceteris paribus. When compared with the main model results, a confusing picture is presented for boys. However, as previously discussed, for the boys sample, the instruments were deemed weaker than the girls and pooled sample; this is particularly so in this model when the additional control is entered into the model. This possible weak instrument problem may therefore assist in producing these results.

For girls, when lagged behaviour scores are controlled for, being placed in the lowest set for maths is found to increase the parent reported externalising behaviours score by 1.7 to 1.8, ceteris paribus. This may be explained possibly by peer effects; since the set level and class behaviour is often correlated, with the behaviour of lower sets being worse (Dunne et al., 2007), children may adopt worse externalising behaviours associated with hyperactivity, attention seeking and obedience when taught with peers who are more likely to exhibit these bad behaviours since children will adopt the behaviours of the roles assigned to them (Campbell, 2013).

The results from this robustness check indicate almost opposite effects to the main models with the behaviour of girls being worsened by setting while boys benefit from setting, though these conclusions are drawn from different models across the analysis. The results in both the main and the robustness analysis are however largely insignificant, thus it could be argued that these results are due to statistical chance.

3.7.3 Excluding children with specific behavioural problems

As an additional check, the IV analysis is undertaken when excluding individuals who are deemed to have specific problems which may influence their behaviour. Children who receive counselling or emotional help at school are excluded from the sample alongside those who receive behavioural support. In addition, children whose SEN is explained by specific behavioural problems are dropped from the sample. Since these children have specific issues relating to their behaviour it is arguable that their behaviour may not be affected in a similar manner to children without specific behavioural and emotional problems. Furthermore, this chapter is interested in identifying the impact of setting for children within the more 'average' range of behavioural issues.

By dropping these children from the sample, 164 observations are lost from the pooled sample analysis; 25% of these children were identified as being in the lowest maths set. The results, do not vary to a great extent with the main results as presented in Table 3.15, though notable differences within the results include an insignificant influence of lowest set placement upon the parental reported internalising scores of boys; in the main results this was a positive and significant effect at the 10% level, thus the effect was initially weakly significant. Once children with greater behavioural issues are dropped from the sample there is no longer a significant influence of lowest maths set placement upon the parental reported behaviours of boys. For girls, a significant negative influence of lowest set placement upon teacher reported internalising behaviours continues to be identified though the results indicate a slightly greater influence of setting upon behaviour; this may be due to children with specific behavioural difficulties being less likely to see changes in behaviour due to a policy such as setting. These results are very briefly presented within Table 3.19 with only the changed results relative to the main model presented. All remaining results maintain the same sign and significance as in the main analysis and are thus unpresented.

3.7.4 High set placement

It is often found within the literature that the low ability and high ability children experience differential impacts of ability grouping with the higher ability children frequently benefitting academically from setting (Hallam and Parsons 2014; Collins and Gan 2013; Ireson 1999a).Just as the previous analysis observed the impact of placement in the lowest maths set upon behaviour, the impact of being placed in the highest maths set will be examined.

Though the IV approach has continued to be adopted throughout the robustness checks with the appropriate IV tests performed each time, the results are presented for the high set placement analysis since this extension encompasses more than simply controlling for an additional variable within the model. The simplified results of the IV tests are provided in Table 3.20; the instruments perform in a similar manner as in the main model where instruments were slightly weaker for boys relative to the girls and the pooled samples. As in the main models, instruments perform generally well though endogeneity is not identified in every model.

The IV results are provided in Tables 3.21 and 3.22 respectively. The OLS results generally indicate that being set and being placed in a low or mid set leads to an increase in behavioural scores relative to not being set; this is given by the set maths coefficient. This gives the opposite effect to that found in the low set analysis since the impact of being placed in a set other than the lowest set produced a negative impact upon behavioural scores; the results from the highest and lowest set analysis therefore seem to be consistent. Thus from these results it seems that lowest set placement generates a greater behavioural score whilst behaviour scores are improved by placement in the higher sets.

The OLS results in Table 3.21 also indicate that being placed in the highest set has a negative effect on behavioural scores, thus, the behaviour of children placed in the highest set for maths improves relative to children who are not set and are taught in mixed ability classes.

The IV results overall provide little evidence of a significant effect of highest set placement upon behaviour. Only for boys is a significant effect found; being placed in the highest set for maths increases the parent reported internalising behaviours score by 1.1, when ability is excluded from the model and by 1.5 when ability is controlled for within the model, ceteris paribus. It therefore seems that for the higher set pupils, setting has little influence upon behaviour. For boys, behaviour actually worsens when placed in the highest set. This finding contrasts with a number of papers within the surrounding literature (Hallam and Parsons 2014; Collins and Gan 2013; Ireson 1999a; Gamoran 2002). The findings suggest that setting generally harms the parent reported internalising behaviours of boys; in both the analysis of this highest and lowest set placement, a positive impact was identified upon these behaviours suggesting that setting may be detrimental to boys' internalising behaviour. This result corresponds somewhat with the findings of Lavy et al. (2012) who identify a negative impact of both 'good' and 'bad' peers, defined as those in the top 5% and the bottom 5% of the ability distribution, upon the performance of boys. As previously argued in the lowest set analysis, internalising behaviours may worsen for boys when learning of their ability level when placed in the lowest set and experiencing a knock in confidence. Based on postulation, it could be argued that for the higher sets, the introduction of more difficult work could decrease boys' confidence since they may possibly struggle more in the higher sets whilst the removal of lower sets removes their reminder of their relative higher ability. It is interesting to note that the negative impact upon internalising behaviours is worse for the lowest set than the highest set boys.

3.8 SUMMARY AND CONCLUSION

This chapter investigated the impact of class setting in mathematics upon the behaviour of primary school children, as measured by the Strengths and Difficulties Questionnaire. Initially, the chapter attempted to identify whether being set for maths influenced the behaviour of children between the ages of 7 and 11; using data from the fourth and fifth waves of the Millennium Cohort Study, an OLS methodology was adopted alongside a Fixed Effects approach in order to overcome the potential problem of unobserved heterogeneity. Additionally, the chapter attempted to estimate the impact of being placed within the lowest maths set upon behaviour, since the literature suggests that the effect of setting may depend upon the ability level or the set level of placement. An OLS and IV methodology was adopted for this part of the analysis in order to overcome the potential endogeneity problem of the child's behaviour and the set placement. The instruments utilised were the proportion of children in the class who spoke English as an additional language and the number of maths sets within the year group of the child. For this part of the analysis, the fifth wave of the MCS was utilised alone, thus the results reflect the impact of being set in maths at primary school when aged 10/11.

The Millennium Cohort Study offers a suitable dataset for this analysis since it provides information on the maths setting experience of respondents in the two waves utilised, the level of set placement for maths and measures of behaviour. Behavioural scores, measured by the total difficulties score alongside internalising and externalising scores, were obtained from both the parent and the teacher of respondents. Since girls and boys are often found to exhibit different types of behaviour, the analysis was undertaken when considering a pooled sample but also when splitting the sample by gender to identify whether setting has a heterogeneous effect according to gender

Results from the fixed effects analysis suggested that the act of setting children for maths in primary school was beneficial for girls' behaviour in terms of the teacher reported internalising behaviour. Since girls are often found to exhibit internalising behaviours whilst boys exhibit externalising behaviours, the findings seemed to support the behavioural literature. In the pooled and boys sample, setting was contrastingly found to be detrimental to behaviour, specifically for teacher reported total difficulties and internalising behaviours. The findings therefore indicate a significant impact of setting in primary school which contrasts with a number of papers within the relevant literature that suggest little influence of setting (Whitburn 2001; Barker Lunn 1970; Kulik and Kulik 1992; Ireson and Hallam, 2005).

A wealth of literature has suggested that higher ability pupils may benefit from setting and other forms of grouping whilst lower ability children are disadvantaged by the practise. This provided the motivation for the second part of the chapter which examined the impact of set placement, specifically in the lowest maths set. Adopting an IV approach, the results indicated that teacher reported internalising behaviours were improved by being placed in the lowest set for maths for girls; it is postulated that this could be due to teaching being tailored to lower ability levels thus preventing girls from becoming withdrawn from the class as in mixed ability class where teaching may be too challenging. This may also improve the confidence of girls' who, within existing literature, are found to suffer from maths anxiety.

Being placed in the lowest set for maths was found to increase the internalising behaviour score of boys as reported by the parent; lowest set placement is therefore found to be detrimental to behaviour. In some models setting was found to be beneficial to the behaviour of alternatively set children; in accordance with the existing literature, these findings identify that the benefits of setting to the higher ability students come at the price of disadvantaging the lower ability pupils (Hallam and Parsons, 2014; Gamoran, 2002; Suknandan and Lee, 1998; Slavin, 1988; Ireson, 1999a). Peer effects may explain the improved behaviour of higher and mid set pupils whilst possibly also explaining the worse behaviour of the low ability pupils.

As a check of robustness, the analysis was also carried out for the highest set pupils; a positive effect of set placement upon the parent reported internalising behaviour scores of boys was again identified, thus both placement in the highest and lowest set had a detrimental impact upon boys' internalising behaviour scores. While the cause of this effect may differ for the low and high sets, the outcome is similar for both boys of high and low ability.

A number of additional robustness checks were carried out including adjustments to the models such as the addition of maternal depression and lagged behavioural score controls; in addition, children with specific behavioural problems were removed from the analysis. When controlling for maternal depression, parent reported internalising behavioural problems continued to be influenced by setting for boys as in the main model, whilst parent reported behaviours remained to be insignificantly influenced by set placement for the pooled sample and for girls. The second robustness check involved dropping children with specific behavioural problems from the analysis; boys' behaviour was insignificantly affected whilst teacher internalising behaviours remained to be significantly influenced for girls. The inclusion of lagged behaviour scores produced some dissimilar results to the main model; externalising behavioural problems were influenced by lowest set placement for both girls and boys whilst internalising behavioural problems were insignificantly impacted by setting. Caution should be taken in the interpretation of the IV models for the boys sample since the instruments proved to be more suitable for the pooled sample and the sample of girls; the instruments suffered from some weakness when adopted in the analysis of the sample of boys. In addition, unlike in all other models, the inclusion of ability altered the results.

This highlights one shortcoming of the chapter since the instruments employed performed very well for the pooled sample and within the sample of girls, they did not work as well when analysing the sample of boys, though it should be noted that the instruments did pass many of the remaining instrumental variable tests. Many instruments were tested in the process yet since providing good instruments for two samples, these were the most suitable. As an alternative approach, future research could adopt other econometric methods, for example propensity score matching, to compare set children to those taught in mixed ability classes. In addition, with data available in future waves of the MCS, it may be interesting to exploit resetting of pupils once they reach secondary school.

The results of this chapter highlight that when estimating the impact of setting or ability grouping it is important to consider the heterogeneous effect by gender. Just as previous research has considered that the impact of setting may not be homogenous across all ability groups, it is clear that this is also true by gender, yet little research in this area has made this consideration. It could also be argued that the reporter of behaviour is important; the results do not indicate that the teacher reported behaviour is consistent with the results of models of parent reported behaviour. Whilst parents are constant and may see only behaviour at home, teachers may be more able to report at-school behaviour though teachers will vary over time for each child. In addition, it is important to consider that behaviour may vary between the home and school environment, possibly explaining differences in reported behaviour.

From a policy perspective, the findings of the chapter may suggest that whilst employing a setting policy within primary schools may assist in improving the behaviour of girls, it may be detrimental to the behaviour of boys both amongst the lower set pupils where behaviour may already be at its poorest, but also amongst highest set boys. This conclusion is, of course, dependent on the model observed and utilised since mixed results are identified across the models; further research is therefore required in order to correctly inform policy. It is likely that there are a number of factors which are influencing the identified effect such as peer effects, specific ability level teaching alongside the attitudes and confidence of pupils which are likely to be influenced by setting.

Table 3.1 Proportion & frequency of sample set for each subject by year

	2008 (WAVE 4 MCS)	2012 (WAVE 5 MCS)
Maths	33.9% Freq: 1812	60.6% Freq: 2406
English/ literature	28.3% Freq: 982	40.7% Freq: 1611

Table 3.2 SDQ questions, sub-scales and measures

SCALE / CATEGORY	ATTRIBUTE	INCLUDED IN TOTAL DIFFICULTIES SCORE	INTERNALISING OR EXTERNALISING BEHAVIOUR
Emotional	Has many worries, often seems worried	✓	Internalising
	Often unhappy, downhearted, tearful	✓	Internalising
	Complains of headache / sickness	✓	Internalising
	Has many fears, is easily scared	✓	Internalising
	Nervous / clingy in new situations	✓	Internalising
Conduct	Often has temper tantrums	✓	Externalising
	Generally obedient	✓	Externalising
	Fights with or bullies other children	✓	Externalising
	Often lies or cheats	✓	Externalising
	Steals from home, school, elsewhere	✓	Externalising
Hyperactivity	Easily distracted, concentration wanders	✓	Externalising
	Sees tasks through to the end	✓	Externalising
	Constantly fidgeting or squirming	✓	Externalising
	Restless, overactive, cannot stay still	✓	Externalising
	Thinks things through before acting	✓	Externalising
Peer	Picked on or bullied by other children	✓	Internalising
	Often solitary, plays alone	✓	Internalising
	Has at least one good friend	✓	Internalising
	Generally liked by other children	✓	Internalising
	Gets on better with adults than children	✓	Internalising
Prosocial	Considerate of other people's feelings		
	Shares readily with other children		
	Helpful if someone is hurt, upset or unwell		
	Kind to younger children		
	Often volunteers to help others		

Source (SDQ info 2014b)

Table 3.3 Sample size summary

	Wave 4 & 5 sample size	**Wave 5 sample size**
Initial	18682	9495
Respondents from Ireland and Scotland dropped	14476	7305
Respondents with missing SDQ data dropped	9800	5106
Respondents with missing or inaccurate EAL data dropped	-	4865

Table 3.4 Variable Definitions

VARIABLE	DESCRIPTION
Teacher response variables	
Set maths	1 if child is set for mathematics, 0 if not.
Low set maths	1 if the child is in the lowest maths set in Y6, 0 otherwise (other set, not set) This is the variable of interest for the second part of the chapter (IV)
Parent interest	1 if the child's mother of father is reported to be very interested or over concerned 0 otherwise.
Mixed year group	1 if child's class contains mixed year group, 0 otherwise.
Class size	Number of children in the child's class.
Teacher tenure	Number of years the teacher has taught at the school.
SEN	1 if the child is classified as having special educational needs, 0 otherwise.
Teacher years	Number of years as a teacher (teacher experience)
Parent Response Variables	
School club	1 if the child attends an after school club, 0 otherwise.
Regular bedtime	1 if the child has a regular bedtime, 0 otherwise.
Ln income	Logged OECD equalised income (income adjusted for size and composition).
Breakfast club	1 if the child attends breakfast club, 0 otherwise.
Parent evening	1 if anyone has attended parents evening, 0 otherwise.
Married	1 if main parent is married or in a civil partnership, 0 otherwise.
Working Household	1 if at least 1 parent from the household is working, 0 otherwise.
Siblings in Household	Number of siblings living in the same household.
Maternal Depression	1 if the Mother is currently treated for depression or anxiety, 0 otherwise.
Time invariant measures	
Maths ability	Maths score given in the wave 4 MCS assessment.
No religious service	1 if the child does not attend religious service, 0 otherwise
Parent degree	1 if one parent has a degree, 0 otherwise.
Male	1 if the child is Male, 0 otherwise.
DOB Autumn/ winter	1 if the child is born in the autumn or winter months, 0 otherwise.
White	1 if the child is white, 0 otherwise.
Birth Order	A continuous variable indicating the birth order of the child.
Birth Weight	A continuous variable indicating the Birth weight of the child in Lbs.
Total SDQ 2008	A continuous variable indicating the 2008 difficulties score. The relevant parent / teacher reported measure is included in the IV analysis.

Table 3.5 Instrumental variable testing

Behavioural outcome measure	(1) Sample	(2) Relevance (P > F)	(3) Over Identification (P - value)	(4) Endogeneity (P - value)	(5) F-statistic	(6) Partial R^2
Total SDQ – Teacher	All	0.00	0.965	0.243	40.307	0.020
	Boys	0.00	0.635	0.573	10.424	0.011
	Girls	0.00	0.572	0.179	30.421	0.030
Total SDQ Parent	All	0.00	0.449	0.803	38.934	0.020
	Boys	0.00	0.432	0.575	9.912	0.010
	Girls	0.00	0.981	0.296	30.617	0.030
Internalising Teacher	All	0.00	0.271	0.095	40.317	0.020
	Boys	0.00	0.663	0.798	10.424	0.011
	Girls	0.00	0.332	0.018	30.478	0.030
Externalising Teacher	All	0.00	0.698	0.693	40.430	0.020
	Boys	0.00	0.502	0.465	10.483	0.011
	Girls	0.00	0.886	0.623	30.542	0.030
Internalising Parent	All	0.00	0.422	0.792	38.84	0.020
	Boys	0.00	0.538	0.032	10.032	0.011
	Girls	0.00	0.971	0.082	30.421	0.030
Externalising Parent	All	0.00	0.581	0.523	38.935	0.020
	Boys	0.00	0.535	0.390	9.912	0.010
	Girls	0.00	0.866	0.934	30.617	0.030

Notes: (2) The null hypothesis of the relevance test: Instrument is uncorrelated with the endogenous regressor
(3) The null hypothesis of the over identification test: all instruments are uncorrelated with the error term (instruments are valid)
(4) Null hypothesis of the endogeneity test: not endogenous (exogenous)
(5) Null hypothesis associated with the F-statistic: Additional instruments have no significant explanatory power. The F statistic should be greater than 10

Table 3.6 Descriptive statistics independent variables (2008 &2012)

Variable		Mean	Std. Dev.	Min	Max
Set Maths	overall	0.499	0.500	0	1
	between		0.453	0	1
	within		0.252	-0.000	0.999
Parent interest	overall	0.921	0.268	0	1
	between		0.263	0	1
	within		0.106	0.421	1.421
Mixed year group	overall	0.246	0.430	0	1
	between		0.406	0	1
	within		0.163	-0.253	0.746
Class size	overall	26.562	4.994	1	64
	between		4.681	1	64
	within		2.017	6.062	47.062
Teacher tenure	overall	8.223	6.975	0	39
	between		6.259	0	39
	within		3.634	-8.776	25.223
SEN	overall	0.118	0.322	0	1
	between		0.307	0	1
	within		0.151	-0.381	0.618
Teacher years	overall	13.227	9.512	0	44
	between		8.476	1	44
	within		4.968	-6.272	32.727
School club	overall	0.299	0.457	0	1
	between		0.419	0	1
	within		0.224	-0.200	0.799
Regular bedtime	overall	0.907	0.289	0	1
	between		0.274	0	1
	within		0.127	0.407	1.407
Ln income	overall	8.521	2.175	2.956	11.235
	between		1.654	2.956	11.235
	within		1.621	4.866	12.176
Breakfast club	overall	0.147	0.355	0	1
	between		0.334	0	1
	within		0.152	-0.352	0.647
Parents evening	overall	0.586	0.492	0	1
	between		0.377	0	1
	within		0.364	0.086	1.086
Married	overall	0.624	0.484	0	1
	between		0.469	0	1
	within		0.149	0.124	1.124
Working household	overall	0.882	0.321	0	1
	between		0.314	0	1

	within		0.109	0.382	1.382
Siblings in HH	overall	1.461	1.017	0	13
	between		1.024	0	13
	within		0.199	-1.038	3.961
Maths ability	overall	9.858	2.753	0	15
	between		2.774	0	15
	within		0	9.858	9.858
Religious service	overall	0.478	0.499	0	1
	between		0.499	0	1
	within		0	0.478	0.478
Parent degree	overall	0.211	0.408	0	1
	between		0.405	0	1
	within		0	0.211	0.211
Male	overall	0.493	0.499	0	1
	between		0.500	0	1
	within		0	0.493	0.493
DOB AW	overall	0.463	0.498	0	1
	between		0.498	0	1
	within		0	0.463	0.463
White	overall	0.851	0.355	0	1
	between		0.366	0	1
	within		0	0.851	0.851
Birth order	overall	0.898	1.001	0	9
	between		1.005	0	9
	within		0	0.898	0.898
Birth Weight	overall	5.577	1.395	0.045	11.530
	between		1.409	0.045	11.530
	within		0	5.577	5.577
T-Bar	1.381				
N	6435				
n	4658				

Notes: The between standard deviation indicates the variation between individuals within the sample whereas the within standard deviation provides the variation within the same individual over time.
'N' provides the number of people for which the overall and within variation is calculated for; this equals the number of observations with non-missing variables
'n' gives the number of individuals. This is the number of people for which the between variation is calculated.
T-bar indicates the average number of time periods / points

Table 3.7 Fixed effects descriptive statistics SDQ measures (2008 & 2012)

Variable		Mean	Std. Dev.	Min	Max
Teacher total difficulties	overall	6.452	5.112	0	35
	between		4.856	0	33
	within		1.884	-7.047	19.952
Parent total difficulties	overall	7.185	5.336	0	34
	between		5.196	0	34
	within		1.556	-4.814	19.185
Teacher internalising	overall	2.440	2.883	0	17
	between		2.679	0	16
	within		1.254	-5.059	9.940
Teacher externalising	overall	2.992	3.490	0	20
	between		3.359	0	19
	within		1.174	-5.007	10.992
Parent internalising	overall	2.875	2.871	0	19
	between		2.757	0	18
	within		1.027	-4.624	10.375
Parent externalising	overall	4.309	3.434	0	20
	between		3.333	0	20
	within		0.966	-0.690	9.309
T-Bar	1.381				
N	6435				
n	4658				

Table 3.8 Setting across years

SET EXPERIENCE	% / FREQUENCY	CHANGE IN SET 2008-2012
Never set (2008 or 2012)	30.47% Freq: 1375	-
Always set (2008 and 2012)	23.18% Freq: 1046	-
Set 2008, not set 2012	9.42% Freq: 425	46.35% Freq: 2091
Not set 2008, set 2012	36.93% Freq: 1666	

Table 3.9 Level of maths set 2012

LEVEL OF MATHS SET	% / FREQUENCY POOLED SAMPLE	% / FREQUENCY GIRLS	% / FREQUENCY BOYS
Lowest set	12% Freq: 488	12.3% Freq: 254	11.6% Freq: 234
Mid Set	17.5% Freq: 715	18.3% Freq: 379	16.7% Freq: 336
Highest Set	31.1% Freq: 1268	29.9% Freq: 618	32.3% Freq: 650
Not Set	39.5% Freq: 1613	39.6% Freq: 819	39.5% Freq: 794
N	4084	2070	2014

Table 3.10 Descriptive statistics of covariates – Level of set analysis

Variable	Mean	Std. Dev.	Min	Max
Set maths	0.605	0.489	0	1
Parent interest	0.906	0.291	0	1
Mixed year group	0.249	0.432	0	1
Class size	26.893	4.978	1	64
Teacher tenure	8.241	6.690	0	39
SEN	0.185	0.389	0	1
Teacher years	13.415	9.062	0	44
School club	0.340	0.474	0	1
Regular bedtime	0.902	0.297	0	1
Ln income	10.203	0.425	8.730	11.235
Breakfast club	0.160	0.367	0	1
Parent evening	0.960	0.196	0	1
Married	0.587	0.492	0	1
Working household	0.881	0.324	0	1
Religious service	0.479	0.500	0	1
Siblings in household	1.472	1.013	0	9
Male	0.493	0.500	0	1
DOB AW	0.452	0.498	0	1
White	0.842	0.364	0	1
Parent degree	0.206	0.405	0	1
Birth order	0.903	0.999	0	6
Birth Weight	5.549	1.417	0.045	11.530
N	4084			
Ability	9.844	2.772	0	15
N	3900			

Notes: These statistics (as in table 11 & 12 also) relate to the analysis of the impact of setting on the level of the set. These statistics therefore predominantly relate to 2012 variables with the exception of DOB, male, birth order and birth weight

Table 3.11 Descriptive statistics of covariates – Lowest set sample

Variable	Mean	Std. Dev.	Min	Max
Set maths	1	0	1	1
Parent interest	0.828	0.378	0	1
Mixed year group	0.160	0.367	0	1
Class size	26.697	4.528	1	40
Teacher tenure	7.794	6.414	0	34
SEN	0.469	0.500	0	1
Teacher years	12.951	9.005	1	44
School club	0.301	0.459	0	1
Regular bedtime	0.869	0.338	0	1
Ln income	10.076	0.419	8.883	10.980
Breakfast club	0.131	0.338	0	1
Parent evening	0.941	0.237	0	1
Married	0.504	0.500	0	1
Working household	0.793	0.406	0	1
Religious service	0.559	0.497	0	1
Siblings in household	1.553	1.016	0	6
Male	0.480	0.500	0	1
DOB AW	0.377	0.485	0	1
White	0.830	0.376	0	1
Parent degree	0.111	0.314	0	1
Birth order	1.070	1.066	0	6
Birth Weight	5.286	1.444	0.045	9.789
N	488			
Ability	7.558	2.845	0	15
N	453			

Table 3.12 Total difficulties summary 2012

SAMPLE	TOTAL DIFFICULTIES MEASURE	MEAN	STD. DEV.	MIN	MAX	N
Full sample Pooled	**Teacher reported**	6.861	4.901	0	33	4084
	Parent reported	7.274	5.515	0	34	4084
Lowest Set Pooled	**Teacher reported**	9.631	5.412	1	31	488
	Parent reported	10.275	6.076	0	32	488
Full sample Boys	**Teacher reported**	7.957	5.444	1	33	2014
	Parent reported	7.857	5.739	0	33	2014
Lowest Set Boys	**Teacher reported**	11.188	5.895	2	31	234
	Parent reported	10.517	6.121	1	32	234
Full sample Girls	**Teacher reported**	5.794	4.033	0	28	2070
	Parent reported	6.708	5.117	0	34	2070
Lowest Set Girls	**Teacher reported**	8.197	4.481	1	23	254
	Parent reported	10.051	6.038	0	31	254

Table 3.13 Does setting influence behaviour? OLS and FE results summary

		OLS	FE	N
Teacher total difficulties score				
	Pooled	0.137	0.317*	6435
		(0.199)	(0.191)	
	Girls	0.050	-0.100	3259
		(0.150)	(0.260)	
	Boys	0.232	0.692**	3176
		(0.185)	(0.281)	
Parent total difficulties score				
	Pooled	0.178	0.038	6435
		(0.126)	(0.160)	
	Girls	0.216	-0.225	3259
		(0.167)	(0.221)	
	Boys	0.147	0.282	3176
		(0.189)	(0.233)	
Teacher reported internalising behaviour				
	Pooled	0.094	0.240*	6435
		(0.072)	(0.129)	
	Girls	-0.009	-0.061	3259
		(0.099)	(0.184)	
	Boys	0.190*	0.523***	3176
		(0.104)	(0.182)	
Teacher reported Externalising behaviour				
	Pooled	0.043	0.055	6435
		(0.080)	(0.119)	
	Girls	0.058	-0.086	3259
		(0.092)	(0.152)	
	Boys	0.048	0.182	3176
		(0.131)	(0.184)	
Parent reported internalising behaviour				
	Pooled	0.023	-0.052	6435
		(0.071)	(0.104)	
	Girls	-0.004	-0.303**	3259
		(0.097)	(0.148)	
	Boys	0.049	0.187	3176
		(0.104)	(0.146)	
Parent reported externalising behaviour				
	Pooled	0.155*	0.090	6435
		(0.081)	(0.100)	
	Girls	0.219**	0.077	3259
		(0.106)	(0.133)	
	Boys	0.098	0.094	3176
		(0.124)	(0.150)	

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table 3.14 Does set placement influence behaviour? – OLS results summary

	OLS (without ability control)			OLS *(controlling for ability)*		
	Pooled	**Girls**	**Boys**	**Pooled**	**Girls**	**Boys**
Teacher total difficulties score						
Lowest set	1.190***	1.264***	1.174***	0.949***	1.070***	0.866**
	(0.228)	(0.268)	(0.374)	(0.238)	(0.279)	(0.394)
Low set	1.749***	1.893***	1.628***	1.474***	1.680***	1.285***
	(0.227)	(0.267)	(0.374)	(0.240)	(0.282)	(0.397)
Set maths	-0.558***	-0.629***	-0.454*	-0.524***	-0.610***	-0.418*
	(0.146)	(0.174)	(0.236)	(0.148)	(0.177)	(0.240)
Parent total difficulties score						
Lowest set	1.422***	2.064***	0.693*	1.242***	1.768***	0.605
	(0.263)	(0.344)	(0.401)	(0.274)	(0.355)	(0.423)
Low set	1.802***	2.634***	0.893**	1.559***	2.299***	0.724*
	(0.262)	(0.343)	(0.400)	(0.276)	(0.358)	(0.427)
Set maths	-0.380**	-0.570**	-0.199	-0.317*	-0.531**	-0.119
	(0.168)	(0.224)	(0.252)	(0.032)	(0.225)	(0.258)
Teacher reported internalising behaviour						
Lowest set	0.611***	0.640***	0.608***	0.502***	0.560***	0.464**
	(0.145)	(0.194)	(0.219)	(0.153)	(0.204)	(0.230)
Low set	0.838***	0.962***	0.731***	0.716***	0.896***	0.551**
	(0.145)	(0.193)	(0.219)	(0.154)	(0.206)	(0.232)
Set maths	-0.227**	-0.321**	-0.122	-0.215**	-0.336***	-0.087
	(0.093)	(0.126)	(0.138)	(0.095)	(0.129)	(0.140)

Teacher reported Externalising behaviour						
Lowest set	0.648***	0.716***	0.609**	0.517***	0.615***	0.435
	(0.153)	(0.159)	(0.266)	(0.160)	(0.165)	(0.280)
Low set	0.984***	1.036***	0.937***	0.833***	0.900***	0.767***
	(0.153)	(0.158)	(0.266)	(0.161)	(0.167)	(0.283)
Set maths	-0.336***	-0.319***	-0.328*	-0.316***	-0.285***	-0.332*
	(0.098)	(0.103)	(0.167)	(0.099)	(0.105)	(0.171)
Parent reported internalising behaviour						
Lowest set	0.561***	0.954***	0.131	0.459***	0.791***	0.069
	(0.151)	(0.205)	(0.221)	(0.158)	(0.214)	(0.233)
Low set	0.745***	1.303***	0.165	0.655***	1.175***	0.090
	(0.150)	(0.205)	(0.220)	(0.159)	(0.216)	(0.235)
Set maths	-0.184*	-0.349***	-0.033	-0.196**	-0.384***	-0.021
	(0.096)	(0.134)	(0.139)	(0.098)	(0.136)	(0.142)
Parent reported externalising behaviour						
Lowest set	0.861***	1.110***	0.567**	0.782***	0.977***	0.536**
	(0.164)	(0.209)	(0.257)	(0.172)	(0.215)	(0.272)
Low set	1.057***	1.331***	0.728***	0.903***	1.124***	0.634**
	(0.164)	(0.209)	(0.256)	(0.173)	(0.217)	(0.274)
Set maths	-0.196*	-0.221	-0.165	-0.121	-0.147	-0.097
	(0.105)	(0.136)	(0.162)	(0.107)	(0.137)	(0.166)
N	4096	2074	2	3911	1993	1918

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table 3.15 Does set placement influence behaviour? IV results summary

	IV (without ability control)			IV *(controlling for ability)*		
	Pooled	**Girls**	**Boys**	**Pooled**	**Girls**	**Boys**
Teacher total difficulties score						
Lowest set	-1.568	-1.451	-1.844	-0.744	-0.589	-0.900
	(1.507)	(1.378)	(3.555)	(1.396)	(1.312)	(3.192)
Low set	-1.721	-1.550	-2.138	-0.644	-0.414	-0.904
	(1.882)	(1.731)	(4.410)	(1.729)	(1.639)	(3.926)
Set maths	0.153	0.099	0.294	-0.101	-0.175	0.004
	(0.407)	(0.402)	(0.893)	(0.368)	(0.378)	(0.776)
Parent total difficulties score						
Lowest set	-0.086	-0.599	1.447	0.858	0.068	2.581
	(1.701)	(1.743)	(3.714)	(1.591)	(1.659)	(3.431)
Low set	-0.087	-0.740	1.838	1.082	0.151	3.171
	(2.123)	(2.191)	(4.609)	(1.972)	(2.073)	(4.219)
Set maths	0.000	0.141	-0.391	-0.224	-0.083	-0.590
	(0.459)	(0.509)	(0.933)	(0.059)	(0.478)	(0.834)
Teacher reported internalising behaviour						
Lowest set	-1.545	-2.294**	-0.564	-1.00	-1.770*	0.137
	(0.974)	(1.041)	(2.050)	(0.903)	(0.991)	(1.851)
Low set	-1.867	-2.757**	-0.728	-1.158	-2.041*	0.148
	(1.217)	(1.308)	(2.543)	(1.119)	(1.238)	(2.277)
Set maths	0.323	0.463	0.164	0.155	0.271	-0.011
	(0.263)	(0.304)	(0.515)	(0.238)	(0.285)	(0.450)

Teacher reported Externalising behaviour						
Lowest set	-0.287	0.595	-1.758	0.060	1.007	-1.411
	(0.992)	(0.786)	(2.540)	(0.930)	(0.767)	(2.295)
Low set	-0.198	0.878	-2.020	0.255	1.390	-1.522
	(1.239)	(0.988)	(3.152)	(1.151)	(0.959)	(2.822)
Set maths	-0.089	-0.283	0.262	-0.195	-0.383*	0.111
	(0.268)	(0.230)	(0.638)	(0.245)	(0.221)	(0.558)
Parent reported internalising behaviour						
Lowest set	0.380	-1.009	3.691*	0.710	-0.798	3.879*
	(0.968)	(1.053)	(2.239)	(0.916)	(1.013)	(2.061)
Low set	0.519	-1.187	4.596*	0.967	-0.833	4.795*
	(1.209)	(1.324)	(2.779)	(1.135)	(1.265)	(2.536)
Set maths	-0.139	0.177	-0.906	-0.257	0.034	-0.915*
	(0.261)	(0.308)	(0.563)	(0.242)	(0.292)	(0.501)
Parent reported externalising behaviour						
Lowest set	-0.467	0.411	-2.244	0.148	0.866	-1.298
	(1.069)	(1.040)	(2.481)	(0.998)	(0.999)	(2.225)
Low set	-0.606	0.447	-2.759	0.116	0.984	-1.623
	(1.335)	(1.307)	(3.078)	(1.237)	(1.248)	(2.736)
Set maths	0.140	-0.036	0.514	0.032	-0.118	0.325
	(0.289)	(0.304)	(0.623)	(0.264)	(0.288)	(0.541)
N	4084	2070	2014	3900	1990	1910

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table 3.16 FE robustness check – inclusion of maternal depression in Parent models. Results summary

	OLS	FE	N
Parent total difficulties score			
Pooled	0.105	0.171	4862
	(0.147)	(0.272)	
Girls	0.134	-0.290	2468
	(0.195)	(0.386)	
Boys	0.073	0.752*	2394
	(0.221)	(0.391)	
Parent reported internalising behaviour			
Pooled	-0.041	-0.011	4862
	(0.084)	(0.177)	
Girls	-0.083	-0.301	2468
	(0.115)	(0.262)	
Boys	-0.005	0.337	2394
	(0.122)	(0.239)	
Parent reported externalising behaviour			
Pooled	0.146	0.182	4862
	(0.094)	(0.170)	
Girls	0.217*	0.011	2468
	(0.122)	(0.222)	
Boys	0.078	0.415	2394
	(0.144)	(0.264)	

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table 3.17 IV robustness checks: including maternal depression. IV results

	IV			IV *(controlling for ability)*		
	Pooled	**Girls**	**Boys**	**Pooled**	**Girls**	**Boys**
Parent total SDQ score						
Lowest set	-0.756	-1.242	0.556	0.241	-0.455	1.686
	(1.706)	(1.761)	(3.613)	(1.579)	(1.649)	(3.311)
Low set	-0.875	-1.489	0.776	0.372	-0.429	2.126
	(2.126)	(2.205)	(4.482)	(1.951)	(2.053)	(4.070)
Set maths	0.119	0.247	-0.220	-0.132	-0.0160	-0.439
	(0.456)	(0.507)	(0.908)	(0.414)	(0.469)	(0.806)
Parent reported internalising behaviour						
Lowest set	0.091	-1.143	2.991	0.410	-0.944	3.255*
	(0.966)	(1.058)	(2.115)	(0.906)	(1.001)	(1.951)
Low set	0.200	-1.316	3.784	0.646	-0.972	4.090*
	(1.204)	(1.325)	(2.623)	(1.120)	(1.251)	(2.399)
Set maths	-0.110	0.173	-0.793	-0.235	0.0284	-0.836*
	(0.258)	(0.305)	(0.531)	(0.237)	(0.286)	(0.475)
Parent reported externalising behaviour						
Lowest set	-0.847	-0.099	-2.436	-0.170	0.498	1,568
	(1.082)	(1.054)	(2.457)	(0.999)	(0.996)	(2.191)
Low set	-1.076	-0.173	-3.008	-0.273	0.543	-1.965
	(1.349)	(1.321)	(3.048)	(1.235)	(1.240)	(2.694)
Set maths	0.229	0.0738	0.573	0.104	-0.0444	0.396
	(0.290)	(0.304)	(0.617)	(0.262)	(0.283)	(0.533)
N	4009	2027	1982	3833	1951	1882

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table 3.18 IV robustness checks: Inclusion of lagged behaviour scores. IV results summary

	IV			IV *(controlling for ability)*		
	Pooled	**Girls**	**Boys**	**Pooled**	**Girls**	**Boys**
Teacher total difficulties score						
Lowest set	-3.129	-0.454	-10.103	-2.659	-0.194	-7.753
	(1.949)	(1.650)	(6.531)	(1.770)	(1.599)	(4.868)
Low set	-3.780	-0.378	-12.52	-3.217	-0.077	-9.616
	(2.434)	(2.068)	(8.115)	(2.211)	(2.003)	(6.054)
Set maths	0.651	-0.075	2.419	0.558	-0.117	1.863
	(0.521)	(0.470)	(1.629)	(0.478)	(0.456)	(1.234)
N	2624	1345	1279	2607	1340	1267
Parent total difficulties score						
Lowest set	0.594	1.460	-1.510	0.696	1.666	-1.523
	(1.363)	(1.385)	(3.046)	(1.275)	(1.333)	(2.701)
Low set	0.774	1.915	-1.844	0.888	2.151	-1.847
	(1.692)	(1.733)	(3.755)	(1.577)	(1.665)	(3.315)
Set maths	-0.180	-0.455	0.334	-0.192	-0.485	0.324
	(0.359)	(0.399)	(0.741)	(0.334)	(0.385)	(0.650)
N	3866	1968	1898	3836	1958	1878
Teacher reported internalising behaviour						
Lowest set	-1.580	-1.135	-3.293	-1.389	-0.977	-2.512
	(1.295)	(1.272)	(3.419)	(1.189)	(1.233)	(2.710)
Low set	-2.016	-1.321	-4.292	-1.778	-1.127	-3.320
	(1.618)	(1.594)	(4.248)	(1.484)	(1.544)	(3.370)
Set maths	0.437	0.186	0.999	0.389	0.150	0.808
	(0.346)	(0.362)	(0.853)	(0.321)	(0.352)	(0.687)

Teacher reported Externalising behaviour						
Lowest set	-1.797	0.614	-7.496	-1.500	0.716	-5.813*
	(1.271)	(0.964)	(4.636)	(1.159)	(0.939)	(3.460)
Low set	-2.102	0.828	-9.108	-1.757	0.934	-7.034
	(1.587)	(1.209)	(5.761)	(1.448)	(1.176)	(4.303)
Set maths	0.305	-0.214	1.611	0.257	-0.218	1.222
	(0.339)	(0.274)	(1.156)	(0.313)	(0.268)	(0.877)
Parent reported internalising behaviour						
Lowest set	0.585	-0.243	2.579	0.575	-0.106	2.113
	(0.874)	(0.931)	(1.968)	(0.818)	(0.893)	(1.711)
Low set	0.822	-0.145	3.233	0.810	0.032	2.650
	(1.085)	(1.165)	(2.427)	(1.012)	(1.115)	(2.100)
Set maths	-0.237	-0.097	-0.654	-0.235	-0.138	-0.536
	(0.230)	(0.268)	(0.479)	(0.215)	(0.258)	(0.412)
Parent reported externalising behaviour						
Lowest set	0.009	1.703*	4.089*	0.121	1.772**	-3.636*
	(0.904)	(0.888)	(2.381)	(0.843)	(0.855)	(2.049)
Low set	-0.048	2.061*	5.077*	0.077	2.119**	-4.497*
	(1.122)	(1.111)	(2.936)	(1.043)	(1.068)	(2.515)
Set maths	0.057	-0.357	0.988*	0.043	-0.347	0.860*
	(0.238)	(0.256)	(0.579)	(0.221)	(0.247)	(0.493)

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table 3.19 IV robustness checks: Removal of children with specified behavioural difficulties

	IV		IV(ability control)	
	Girls	**Boys**	**Girls**	**Boys**
Teacher internalising				
Lowest set	-2.496**	-	-1.961**	-
	(1.083)		(1.026)	
Low set	-3.019**		-2.290*	
	(1.355)		(1.275)	
Set maths	0.523*		0.329	
	(0.308)		(0.287)	
N	2,043		1,964	
Parent internalising				
Lowest set	-	3.164	-	3.362
		(2.275)		(2.162)
Low set		3.825		4.038
		(2.774)		(2.613)
Set maths		-0.661		-0.676
		(0.522)		(0.477)
N		1910		1818

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table 3.20 IV tests: Instrumenting higher set placement

Behavioural outcome measure	(1) Sample	(2) Over Identification (P - value)	(3) Endogeneity (P - value)	(4) F-statistic
Total SDQ – Teacher	All	0.317	0.377	22.358
	Boys	0.413	0.732	13.646
	Girls	0.598	0.386	9.875
Total SDQ Parent	All	0.616	0.169	24.049
	Boys	0.455	0.357	14.967
	Girls	0.937	0.448	10.239

Table 3.21 Placement in highest maths set – OLS results

	OLS (without ability control)			OLS *(controlling for ability)*		
	Pooled	**Girls**	**Boys**	**Pooled**	**Girls**	**Boys**
Teacher total difficulties score						
Highest set	-0.957***	-1.046***	-0.841**	-0.859***	-0.941***	-0.767***
	(0.164)	(0.196)	(0.295)	(0.168)	(0.202)	(0.271)
High set	-1.539***	-1.612***	-1.494***	-1.287***	-1.348***	-1.265***
	(0.179)	(0.211)	(0.293)	(0.191)	(0.224)	(0.314)
Set maths	0.582***	0.566***	0.652**	0.429**	0.408**	0.498*
	(0.165)	(0.195)	(0.271)	(0.171)	(0.201)	(0.282)
Parent total difficulties score						
Highest set	-0.822***	-1.183***	-0.471*	-0.589***	-0.983***	-0.206
	(0.189)	(0.251)	(0.282)	(0.194)	(0.256)	(0.291)
High set	-1.649***	-2.309***	0.944***	-1.193***	-1.843***	-0.479
	(0.206)	(0.270)	(0.314)	(0.220)	(0.285)	(0.339)
Set maths	0.827***	1.126***	0.473	0.604	0.860***	0.274
	(0.191)	(0.250)	(0.291)	(0.197)	(0.255)	(0.304)
Teacher reported internalising behaviour						
Highest set	-0.365***	-0.546***	-0.187	-0.310***	-0.515***	-0.114
	(0.105)	(0.142)	(0.154)	(0.109)	(0.147)	(0.158)
High set	-0.627***	-0.845***	-0.439**	-0.488***	-0.726***	-0.281
	(0.114)	(0.152)	(0.172)	(0.120)	(0.164)	(0.184)
Set maths	0.263**	0.299**	0.252	0.178	0.210	0.167
	(0.106)	(0.141)	(0.159)	(0.111)	(0.147)	(0.165)

Teacher reported Externalising behaviour						
Highest set	-0.598***	-0.516***	0.649***	-0.560***	-0.446***	0.656***
	(0.109)	(0.117)	(0.187)	(0.113)	(0.120)	(0.193)
High set	-0.943***	-0.819***	1.066***	0.840***	-0.690***	1.000***
	(0.120)	(0.125)	(0.208)	(0.128)	(0.133)	(0.224)
Set maths	0.345***	0.304***	0.417**	0.280**	0.244**	0.345*
	(0.111)	(0.116)	(0.193)	(0.115)	(0.119)	(0.201)
Parent reported internalising behaviour						
Highest set	-0.287***	-0.531***	-0.062	-0.207*	-0.489***	0.064
	(0.108)	(0.151)	(0.155)	(0.111)	(0.155)	(0.161)
High set	-0.519***	-0.902***	-0.128	-0.289**	-0.693***	0.145
	(0.118)	(0.162)	(0.173)	(0.127)	(0.172)	(0.186)
Set maths	0.232**	0.371**	0.0662	0.0821	0.203	-0.0807
	(0.110)	(0.150)	(0.160)	(0.114)	(0.154)	(0.167)
Parent reported externalising behaviour						
Highest set	-0.535***	-0.652***	-0.409**	-0.382***	-0.493***	-0.270
	(0.118)	(0.152)	(0.180)	(0.121)	(0.155)	(0.188)
High set	-1.130***	-1.407***	0.816***	-0.904***	-1.151***	0.624***
	(0.129)	(0.164)	(0.201)	(0.137)	(0.172)	(0.218)
Set maths	0.595***	0.755***	0.407**	0.522***	0.657***	0.354*
	(0.119)	(0.151)	(0.186)	(0.123)	(0.154)	(0.195)
N	4096	2074	2022	3911	1993	1918

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table 3.22 Placement in highest maths set – IV results

	IV (without ability control)			IV *(controlling for ability)*		
	Pooled	**Girls**	**Boys**	**Pooled**	**Girls**	**Boys**
Teacher total difficulties score						
Highest set	-0.549	-0.463	-0.633	-0.208	-0.179	-0.380
	(0.774)	(1.091)	(0.932)	(0.850)	(1.129)	(1.194)
High set	-0.722	-0.466	-1.076	0.0402	0.163	-0.457
	(1.555)	(2.134)	(2.174)	(1.728)	(2.225)	(2.507)
Set maths	0.173	0.00342	0.442	-0.248	-0.342	0.0763
	(0.806)	(1.068)	(1.167)	(0.901)	(1.120)	(1.354)
Parent total difficulties score						
Highest set	0.134	-0.354	0.410	0.761	0.099	1.018
	(0.900)	(1.403)	(1.134)	(0.994)	(1.440)	(1.302)
High set	0.307	-0.672	0.934	1.591	0.301	2.159
	(1.808)	(2.745)	(2.344)	(2.020)	(2.835)	(2.733)
Set maths	-0.174	0.318	-0.524	-0.831	-0.202	-1.140
	(0.936)	(1.374)	(1.259)	(1.053)	(1.428)	(1.476)
Teacher reported internalising behaviour						
Highest set	-0.617	-1.087	-0.153	-0.389	-0.845	0.080
	(0.495)	(0.792)	(0.616)	(0.542)	(0.817)	(0.699)
High set	-1.147	-1.918	-0.368	-0.652	-1.384	0.135
	(0.995)	(1.550)	(1.274)	(1.104)	(1.610)	(1.467)
Set maths	0.530	0.831	0.215	0.263	0.539	-0.0553
	(0.515)	(0.776)	(0.684)	(0.575)	(0.811)	(0.792)

Teacher reported Externalising behaviour						
Highest set	-0.060	0.532	-0.614	0.072	0.646	-0.600
	(0.524)	(0.686)	(0.746)	(0.574)	(0.705)	(0.848)
High set	0.141	1.248	-1.017	0.446	1.478	-0.907
	(1.052)	(1.341)	(1.542)	(1.167)	(1.389)	(1.780)
Set maths	-0.201	-0.715	0.403	-0.375	-0.832	0.306
	(0.545)	(0.671)	(0.828)	(0.608)	(0.700)	(0.961)
Parent reported internalising behaviour						
Highest set	0.280	-0.728	1.094*	0.521	-0.631	1.459**
	(0.517)	(0.837)	(0.649)	(0.571)	(0.859)	(0.751)
High set	0.639	-1.296	2.327*	1.212	-0.984	3.140**
	(1.039)	(1.636)	(1.341)	(1.161)	(1.692)	(1.577)
Set maths	-0.360	0.568	-1.233*	-0.691	0.353	1.681**
	(0.538)	(0.819)	(0.720)	(0.605)	(0.852)	(0.851)
Parent reported externalising behaviour						
Highest set	-0.146	0.363	-0.684	0.240	0.730	-0.440
	(0.557)	(0.873)	(0.722)	(0.615)	(0.901)	(0.824)
High set	-0.332	0.624	-1.393	0.380	1.284	-0.982
	(1.120)	(1.707)	(1.492)	(1.250)	(1.774)	(1.731)
Set maths	0.186	-0.251	0.709	-0.140	-0.554	0.541
	(0.580)	(0.854)	(0.801)	(0.652)	(0.893)	(0.935)
N	4,084	2070	2014	3900	1990	1910

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

3.9 APPENDIX

Table A3.1 Does setting influence behaviour? Full results pooled sample internalising & externalising behaviours

VARIABLES	OLS Teacher total difficulties	FE Teacher total difficulties	OLS Parent total difficulties	FE Parent total difficulties
Set maths	0.137	0.317*	0.178	0.038
	(0.119)	(0.191)	(0.126)	(0.160)
Parental interest	-2.962***	-1.180***	-1.412***	-0.007
	(0.222)	(0.416)	(0.236)	(0.349)
Mixed year group	0.128	-0.235	-0.066	-0.248
	(0.133)	(0.271)	(0.141)	(0.227)
Class size	-0.017	-0.044**	-0.005	-0.009
	(0.011)	(0.022)	(0.012)	(0.018)
Teacher tenure	-0.038***	-0.028	-0.002	0.007
	(0.012)	(0.017)	(0.012)	(0.014)
SEN	3.239***	0.710**	3.656***	1.454***
	(0.190)	(0.317)	(0.202)	(0.266)
Teaching years	0.014	-0.007	0.008	-0.008
	(0.008)	(0.012)	(0.009)	(0.010)
School club	-0.010	-0.226	0.073	0.049
	(0.127)	(0.199)	(0.135)	(0.167)
Regular bedtime	-0.274	0.318	-1.200***	0.106
	(0.198)	(0.345)	(0.210)	(0.289)
Ln income	-0.056	-0.020	-0.324***	0.038
	(0.077)	(0.122)	(0.082)	(0.103)
Breakfast club	0.563***	0.179	0.481***	0.153
	(0.163)	(0.291)	(0.173)	(0.244)
Parents evening	0.365	0.869	0.648*	-0.160
	(0.337)	(0.539)	(0.358)	(0.452)
Married	-0.667***	-0.149	-0.959***	-0.409
	(0.127)	(0.309)	(0.135)	(0.259)
Working household	-1.136***	-0.456	-1.063***	-0.173
	(0.198)	(0.409)	(0.210)	(0.343)
Siblings in HH	-0.074	0.030	0.211***	-0.175
	(0.071)	(0.223)	(0.075)	(0.187)
ability	-0.299***		-0.309***	
	(0.022)		(0.023)	
Degree	-0.663***		-1.032***	
	(0.148)		(0.157)	
Male	1.695***		0.962***	
	(0.116)		(0.123)	
DOB A/W	-0.439***		-0.351***	
	(0.114)		(0.121)	
White	0.354**		0.002	
	(0.172)		(0.183)	
Birth order	-0.022		-0.373***	
	(0.071)		(0.075)	
Birth Weight lbs	-0.098**		-0.097**	
	(0.042)		(0.044)	
No religious service	-0.013		0.167	
	(0.120)		(0.128)	
Constant	13.98***	8.742***	16.39***	7.660***
	(0.679)	(1.167)	(0.720)	(0.978)
Observations	6,435	6,435	6,435	6,435
R-squared	0.212	0.055	0.185	0.026

Standard errors in parentheses *** p<001, ** p<0.05, * p<0.1

Table A3.2 Does setting influence behaviour? Full results pooled sample internalising & externalising behaviours

VARIABLES	FE Teacher internalising	FE Teacher externalising	FE Parent internalising	FE Parent externalising
Set maths	0.240*	0.055	-0.052	0.090
	(0.129)	(0.119)	(0.104)	(0.100)
Parental interest	-0.460	-0.855***	-0.041	0.033
	(0.282)	(0.259)	(0.226)	(0.217)
Mixed year group	0.016	-0.211	-0.174	-0.075
	(0.184)	(0.168)	(0.147)	(0.141)
Class size	-0.030**	-0.016	-0.004	-0.006
	(0.015)	(0.014)	(0.012)	(0.012)
Teacher tenure	-0.011	-0.018*	0.003	0.005
	(0.012)	(0.011)	(0.010)	(0.009)
SEN	0.762***	0.193	0.939***	0.515***
	(0.215)	(0.197)	(0.173)	(0.166)
teaching years	-0.001	-0.008	-0.006	-0.002
	(0.009)	(0.008)	(0.007)	(0.007)
School club	-0.251*	0.009	0.004	0.046
	(0.135)	(0.124)	(0.108)	(0.104)
Regular bedtime	0.118	0.259	0.280	-0.175
	(0.234)	(0.214)	(0.187)	(0.180)
Ln income	0.001	-0.303***	0.075	-0.036
	(0.083)	(0.076)	(0.066)	(0.064)
Breakfast club	0.004	0.137	0.191	-0.037
	(0.197)	(0.181)	(0.158)	(0.152)
Parents evening	-0.272	0.641*	0.079	-0.239
	(0.366)	(0.335)	(0.293)	(0.281)
Married	-0.019	-0.130	-0.196	-0.214
	(0.209)	(0.192)	(0.168)	(0.161)
Working household	-0.383	0.044	0.019	-0.192
	(0.277)	(0.254)	(0.222)	(0.213)
Siblings in HH	-0.091	0.103	-0.004	-0.171
	(0.151)	(0.139)	(0.121)	(0.117)
Constant	4.143***	6.287***	2.184***	5.476***
	(0.791)	(0.726)	(0.634)	(0.609)
Observations	6,435	6,435	6,435	6,435
R-squared	0.019	0.059	0.061	0.020

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A3.3 Does set placement influence behaviour? Pooled sample total difficulties

VARIABLES	IV Teacher total difficulties	IV Teacher total difficulties	IV Parent total difficulties	IV Parent total difficulties
Lowest set	-1.568	-0.744	-0.086	0.858
	(1.507)	(1.396)	(1.701)	(1.591)
Low set	-1.721	-0.644	-0.087	1.082
	(1.882)	(1.729)	(2.123)	(1.972)
Set maths	0.153	-0.101	0.000	-0.224
	(0.407)	(0.368)	(0.459)	(0.420)
Ability		-0.220***		-0.225***
		(0.052)		(0.059)
Parental interest	-3.066***	-2.851***	-1.312***	-1.156***
	(0.273)	(0.261)	(0.308)	(0.298)
Mixed year group	0.096	0.149	-0.040	-0.036
	(0.178)	(0.175)	(0.201)	(0.200)
Class size	-0.018	-0.009	-0.019	-0.006
	(0.014)	(0.014)	(0.015)	(0.016)
Teacher tenure	-0.031**	-0.027*	0.005	0.002
	(0.014)	(0.014)	(0.016)	(0.016)
SEN	4.186***	3.517***	4.221***	3.542***
	(0.440)	(0.337)	(0.496)	(0.385)
Teaching years	0.007	0.008	0.000	0.006
	(0.010)	(0.010)	(0.012)	(0.012)
School club	-0.023	-0.008	-0.082	-0.075
	(0.148)	(0.147)	(0.167)	(0.168)
Regular bedtime	-0.233	-0.324	-0.983***	-0.994***
	(0.233)	(0.232)	(0.263)	(0.265)
Ln income	-0.824***	-0.541**	-1.867***	-1.620***
	(0.261)	(0.256)	(0.295)	(0.292)
Breakfast club	0.439**	0.476**	0.304	0.286
	(0.192)	(0.192)	(0.216)	(0.219)
Parents evening	-0.285	-0.197	-0.348	-0.252
	(0.359)	(0.358)	(0.405)	(0.408)
Married	-0.507***	-0.490***	-0.734***	-0.720***
	(0.152)	(0.151)	(0.171)	(0.172)
Working household	-0.708***	-0.788***	-0.722**	-0.738**
	(0.252)	(0.258)	(0.284)	(0.294)
No religious service	0.129	0.087	0.114	0.035
	(0.151)	(0.148)	(0.170)	(0.169)
Siblings in HH	-0.221**	-0.136	-0.195*	-0.116
	(0.092)	(0.092)	(0.105)	(0.106)
Male	1.569***	1.623***	0.565***	0.719***
	(0.153)	(0.147)	(0.173)	(0.168)
DOB A/W	-0.354**	-0.299**	-0.377**	-0.337**
	(0.149)	(0.149)	(0.168)	(0.170)
White	0.490**	0.544**	0.223	0.218
	(0.220)	(0.218)	(0.248)	(0.249)
Degree	-0.662***	-0.526***	-0.647***	-0.535**
	(0.202)	(0.196)	(0.228)	(0.223)
Birth order	0.091	0.011	-0.172*	-0.275***
	(0.084)	(0.083)	(0.095)	(0.094)
Birth Weight lbs	-0.079	-0.054	-0.040	0.007
	(0.052)	(0.050)	(0.058)	(0.057)
Constant	19.07***	17.77***	29.93***	28.76***
	(2.672)	(2.667)	(3.015)	(3.041)
Observations	4,084	3,900	4,084	3,900
R-squared	0.206	0.239	0.187	0.207

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table A3.4 Does set placement influence behaviour? Pooled sample teacher & parent reported difficulties

VARIABLES	IV teacher internalising	IV teacher internalising	IV teacher externalising	IV teacher externalising	IV Parent internalising	IV Parent internalising	IV Parent externalising	IV Parent externalising
Lowest set	-1.545	-1.00	-0.287	0.060	0.380	0.710	-0.467	0.148
	(0.974)	(0.903)	(0.992)	(0.930)	(0.968)	(0.916)	(1.069)	(0.998)
Low set	-1.867	-1.158	-0.198	0.255	0.519	0.967	-0.606	0.116
	(1.217)	(1.119)	(1.239)	(1.151)	(1.209)	(1.135)	(1.335)	(1.237)
Set maths	0.323	0.155	-0.089	-0.195	-0.139	-0.257	0.140	0.032
	(0.263)	(0.238)	(0.268)	(0.245)	(0.261)	(0.242)	(0.289)	(0.264)
Ability		-0.122***		-0.107***		-0.081**		-0.144***
		(0.033)		(0.034)		(0.034)		(0.037)
Parental interest	-1.147***	-1.045***	-2.174***	-2.034***	-0.277	-0.221	-1.035***	-0.934***
	(0.177)	(0.169)	(0.180)	(0.174)	(0.176)	(0.171)	(0.194)	(0.187)
Mixed year group	0.031	0.060	0.070	0.095	0.031	0.030	-0.071	-0.066
	(0.115)	(0.113)	(0.117)	(0.117)	(0.114)	(0.115)	(0.126)	(0.125)
Class size	-0.001	-0.000	-0.018**	-0.010	0.000	0.005	-0.019**	-0.012
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)
Teacher tenure	-0.014	-0.012	-0.019*	-0.017*	0.008	0.008	-0.003	-0.005
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)
SEN	2.235***	1.865***	2.253***	1.919***	1.898***	1.660***	2.322***	1.882***
	(0.284)	(0.218)	(0.289)	(0.225)	(0.282)	(0.221)	(0.312)	(0.241)
Teaching years	0.006	0.005	0.001	0.003	-0.005	-0.003	0.006	0.010
	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
School club	-0.078	-0.068	0.067	0.074	0.044	0.033	-0.127	-0.109
	(0.095)	(0.095)	(0.097)	(0.098)	(0.095)	(0.096)	(0.105)	(0.105)
Regular bedtime	-0.007	-0.036	-0.231	-0.299*	-0.132	-0.125	-0.850***	-0.869***
	(0.151)	(0.150)	(0.153)	(0.154)	(0.150)	(0.152)	(0.165)	(0.166)
Ln income	-0.207	-0.078	-0.674***	-0.511***	-0.705***	-0.650***	-1.161***	-0.969***
	(0.169)	(0.166)	(0.172)	(0.170)	(0.168)	(0.168)	(0.185)	(0.183)
Breakfast club	0.170	0.211*	0.320**	0.326**	-0.048	-0.040	0.352***	0.327**
	(0.124)	(0.125)	(0.126)	(0.128)	(0.123)	(0.126)	(0.136)	(0.138)
Parents evening	-0.271	-0.260	-0.149	-0.055	0.068	0.051	-0.416	-0.303
	(0.232)	(0.232)	(0.237)	(0.238)	(0.231)	(0.235)	(0.255)	(0.256)
Married	-0.230**	-0.227**	-0.331***	-0.313***	-0.380***	-0.399***	-0.354***	-0.321***
	(0.098)	(0.097)	(0.099)	(0.100)	(0.097)	(0.098)	(0.108)	(0.108)

Working household	-0.541***	-0.527***	-0.198	-0.302*	-0.386**	-0.357**	-0.336*	-0.380**
	(0.163)	(0.167)	(0.166)	(0.172)	(0.162)	(0.169)	(0.179)	(0.185)
No religious service	0.015	-3.80e-05	0.110	0.082	-0.062	-0.076	0.176*	0.112
	(0.097)	(0.096)	(0.099)	(0.098)	(0.096)	(0.097)	(0.107)	(0.106)
Siblings in HH	-0.129**	-0.081	-0.102*	-0.061	-0.098*	-0.062	-0.096	-0.053
	(0.060)	(0.059)	(0.061)	(0.061)	(0.059)	(0.060)	(0.065)	(0.066)
Male	-0.112	-0.071	1.813***	1.833***	-0.248**	-0.201**	0.812***	0.920***
	(0.099)	(0.095)	(0.101)	(0.097)	(0.098)	(0.096)	(0.109)	(0.105)
DOB A/W	-0.090	-0.073	-0.271***	-0.228**	0.012	0.024	-0.389***	-0.361***
	(0.096)	(0.096)	(0.097)	(0.099)	(0.095)	(0.098)	(0.105)	(0.107)
White	0.376***	0.403***	0.107	0.134	-0.158	-0.132	0.381**	0.350**
	(0.142)	(0.141)	(0.145)	(0.145)	(0.141)	(0.143)	(0.156)	(0.156)
Degree	-0.381***	-0.300**	-0.295**	-0.234*	-0.112	-0.070	-0.535***	-0.465***
	(0.131)	(0.127)	(0.133)	(0.130)	(0.130)	(0.128)	(0.143)	(0.140)
Birth order	0.038	-0.009	0.050	0.014	-0.099*	-0.148***	-0.071	-0.126**
	(0.054)	(0.053)	(0.055)	(0.055)	(0.054)	(0.054)	(0.059)	(0.059)
Birth Weight lbs	-0.075**	-0.061*	-0.010	-0.000	0.005	0.023	-0.046	-0.015
	(0.033)	(0.032)	(0.034)	(0.033)	(0.033)	(0.033)	(0.036)	(0.036)
Constant	6.551***	6.219***	11.84***	10.82***	11.29***	11.22***	18.64***	17.54***
	(1.727)	(1.725)	(1.759)	(1.776)	(1.716)	(1.750)	(1.895)	(1.907)
Observations	4,084	3,900	4,084	3,900	4,084	3,900	4,084	3,900
R-squared	0.040	0.088	0.263	0.271	0.112	0.120	0.177	0.200

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

Table A3.5 Does set placement influence behaviour? Total difficulties by gender

VARIABLES	Girls IV Teacher total difficulties	Girls IV Teacher total difficulties	Girls IV Parent total difficulties	Girls IV Parent total difficulties	Boys IV Teacher total difficulties	Boys IV Teacher total difficulties	Boys IV Parent total difficulties	Boys IV Parent total difficulties
Lowest set	-1.451	-0.589	-0.599	0.068	-1.844	-0.900	1.447	2.581
	(1.378)	(1.312)	(1.743)	(1.659)	(3.555)	(3.192)	(3.714)	(3.431)
Low set	-1.550	-0.414	-0.740	0.151	-2.138	-0.904	1.838	3.171
	(1.731)	(1.639)	(2.191)	(2.073)	(4.410)	(3.926)	(4.609)	(4.219)
Set maths	0.099	-0.175	0.141	-0.083	0.294	0.004	-0.391	-0.590
	(0.402)	(0.378)	(0.509)	(0.478)	(0.893)	(0.776)	(0.933)	(0.834)
Ability		-0.211***		-0.315***		-0.228**		-0.129
		(0.060)		(0.076)		(0.094)		(0.101)
Parental interest	-3.287***	-3.215***	-1.441***	-1.297***	-2.855***	-2.497***	-1.162**	-1.055**
	(0.339)	(0.331)	(0.429)	(0.419)	(0.444)	(0.402)	(0.464)	(0.432)
Mixed year group	0.232	0.277	-0.168	-0.251	-0.077	-0.005	0.132	0.220
	(0.207)	(0.204)	(0.262)	(0.257)	(0.306)	(0.297)	(0.319)	(0.319)
Class size	0.024	0.034*	0.024	0.042*	-0.057***	-0.051**	-0.065***	-0.059**
	(0.017)	(0.018)	(0.022)	(0.022)	(0.021)	(0.022)	(0.022)	(0.023)
Teacher tenure	-0.041**	-0.035**	-0.033	-0.036*	-0.014	-0.013	0.043*	0.044*
	(0.017)	(0.017)	(0.022)	(0.022)	(0.023)	(0.023)	(0.024)	(0.025)
SEN	3.699***	3.080***	3.981***	3.268***	4.522***	3.786***	4.066***	3.491***
	(0.460)	(0.372)	(0.582)	(0.471)	(0.955)	(0.702)	(0.998)	(0.754)
Teaching years	0.017	0.018	0.016	0.021	-0.003	-0.002	-0.014	-0.008
	(0.013)	(0.013)	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	(0.018)
School club	-0.037	0.020	-0.311	-0.215	0.073	0.059	0.132	0.056
	(0.180)	(0.178)	(0.228)	(0.225)	(0.239)	(0.240)	(0.250)	(0.258)
Regular bedtime	-0.323	-0.334	-0.828**	-0.669*	-0.091	-0.263	-1.147***	-1.308***
	(0.288)	(0.284)	(0.365)	(0.360)	(0.371)	(0.367)	(0.387)	(0.394)

Log Income	-0.813**	-0.509	-2.148***	-1.698***	-0.825**	-0.548	-1.602***	-1.470***
	(0.327)	(0.315)	(0.414)	(0.398)	(0.405)	(0.403)	(0.423)	(0.433)
Breakfast club	0.461**	0.467**	0.404	0.356	0.383	0.449	0.222	0.225
	(0.233)	(0.231)	(0.294)	(0.292)	(0.305)	(0.309)	(0.319)	(0.332)
Parents evening	0.270	0.333	-0.722	-0.646	-0.751	-0.612	-0.005	0.120
	(0.456)	(0.447)	(0.576)	(0.565)	(0.551)	(0.554)	(0.575)	(0.596)
Married	-0.347*	-0.348*	-0.589**	-0.519**	-0.696***	-0.667***	-0.876***	-0.929***
	(0.185)	(0.182)	(0.235)	(0.230)	(0.239)	(0.240)	(0.250)	(0.258)
Working household	-0.766**	-0.658**	-0.831**	-0.696*	-0.657	-0.946**	-0.630	-0.829*
	(0.306)	(0.310)	(0.387)	(0.392)	(0.405)	(0.418)	(0.424)	(0.449)
No religious service	0.020	-0.098	0.241	0.051	0.242	0.294	0.021	0.028
	(0.183)	(0.176)	(0.232)	(0.223)	(0.241)	(0.242)	(0.251)	(0.260)
Siblings in HH	-0.230**	-0.137	-0.398***	-0.313**	-0.225	-0.137	-0.002	0.083
	(0.116)	(0.115)	(0.147)	(0.145)	(0.146)	(0.146)	(0.152)	(0.157)
DOB A/W	0.001	-0.007	-0.288	-0.318	-0.729**	-0.620**	-0.398	-0.278
	(0.169)	(0.168)	(0.214)	(0.212)	(0.286)	(0.290)	(0.299)	(0.312)
White	0.615**	0.681***	0.367	0.295	0.323	0.339	0.050	0.117
	(0.262)	(0.261)	(0.332)	(0.330)	(0.363)	(0.355)	(0.379)	(0.381)
Degree	-0.266	-0.215	-0.669**	-0.630**	-1.118***	-0.889***	-0.489	-0.346
	(0.237)	(0.235)	(0.300)	(0.297)	(0.388)	(0.340)	(0.406)	(0.366)
Birth order	0.120	0.062	0.003	-0.055	0.097	-0.019	-0.369**	-0.511***
	(0.100)	(0.100)	(0.127)	(0.127)	(0.142)	(0.134)	(0.149)	(0.144)
Birth Weight lbs	-0.040	-0.019	-0.056	0.011	-0.129	-0.099	-0.012	0.010
	(0.061)	(0.060)	(0.077)	(0.076)	(0.092)	(0.084)	(0.096)	(0.090)
Constant	16.98***	15.29***	32.22***	29.44***	22.45***	21.43***	28.47***	28.16***
	(3.372)	(3.255)	(4.266)	(4.117)	(4.079)	(4.257)	(4.262)	(4.575)
Observations	2,070	1,990	2,070	1,990	2,014	1,910	2,014	1,910
R-squared	0.142	0.194	0.146	0.185	0.187	0.218	0.201	0.198

*** p<001, ** p<0.05, * p<0.1 . Standard errors in parentheses

CHAPTER 4 : THE EFFECT OF PRIMARY CONVERTER ACADEMIES ON PUPIL PERFORMANCE

4.1 INTRODUCTION

Introduced by the Labour government in the early 2000s, academy schools in England represent a step away from the traditional model of local-authority controlled schooling. Initially, the academy programme aimed to raise standards in education by targeting underperforming schools, predominantly in deprived areas and converting these schools into sponsored academies. Such schools would be managed by an academy trust and would be appointed a new governing body. By 2010, 203 secondary sponsored academies existed within England's education system (Department for Education, 2011).

In 2010, the coalition government took action to expand the existing academies programme; not only were underperforming primary schools also targeted to become sponsored academies, but converter academies[17] were introduced into all levels of the education system, since no longer confined to secondary schools. The academies programme was opened to all primary and secondary schools in England thus allowing for schools to apply to voluntarily convert to become academies to benefit from the greater autonomy that academy status offered to schools.

Though the proportion of primary academies is relatively small at 13% relative to the 60% of secondary schools that academies constitute, the number of primary academies now exceeds the number of secondary academies (House of commons, 2015) with over half a million pupils attending a primary academy (Department for Education, 2014c). Furthermore, the 2016 budget set out plans for all schools to become academies by 2020; although this blanket policy applying to all schools was reconsidered, the Conservative government have expressed their continued commitment to expanding the academy programme (Department for Education, 2016).

[17] Sponsored academies are usually previously underperforming schools that convert under a sponsor to become an academy in order to raise attainment. Converter academies are previously successful schools that voluntarily select to become an academy, often to benefit from greater autonomy.

Few papers have adopted econometric techniques to overcome the evaluation problem[18] in order to analyse the impact of converter academies since the policy overhaul in 2010. The existing literature largely focuses on the impact of academies prior to 2010, or attends to the analysis of sponsored academies. The sponsored academy literature provides mixed results on the influence upon pupil outcomes; whilst a number of papers identify a beneficial impact upon pupil outcomes (Department for Education 2012; Hutchings et al. 2014), the work of Wilson (2011) suggests that sponsored academies are failing disadvantaged pupils. Of the few papers that do investigate converter academies, the results seem promising when considering the impact upon pupil outcomes at the secondary school level (Eyles and Machin, 2015).

At present, there is very little evidence of the impact of academy status on pupil outcome at primary school level, thus prompting the reported need for research into this matter by the House of Commons (2015). Furthermore, little research has been undertaken to examine the performance of post-2010 converter academies.

Though it is important to evaluate the academies policy, it is imperative to understand the impact of the academy programme upon the progress of primary school children. Primary education is fundamental to an individual's development and lifelong learning while educational attainment in the early stages of schooling are key determinants of educational outcomes later in life (Dearden et al. 2004). Evidence also suggests that interventions at the earliest stages of education may counteract the negative influences of poor family background (Heckman, 2000). Since one in five children leave primary school unable to read at the level required for secondary school (Department for Education, 2015b), it is vital to understand which policies do improve educational standards in order to use resources in the most efficient manner.

This chapter will utilise individual-level data to analyse the impact of primary converter academies upon pupil cognitive outcomes, specifically focusing on the impact upon the percentile rank of pupils according to their average point score (APS). The analysis will employ a research design that exploits the differential timing of academy conversion between primary schools, for those schools that did voluntarily convert; this allows for a credible control group to be identified for the purposes of the analysis.

[18] See chapter 2 for a full discussion of the evaluation problem and section 4.5 of this chapter

Specifically, the chapter will adopt a difference-in-differences (DID) methodology to analyse the impact of converter academies on pupils who enrolled in the primary school before conversion. The outcomes of these treated individuals will be compared with a control group of pupils who attended schools that eventually became converter academies; the schools of the control group became academies once the pupil had left the primary school.

This approach addresses a number of potential issues; firstly, research within the academies literature suggests that pupil intake may vary once a school becomes an academy (Eyles and Machin, 2015; Wilson, 2011); academies may then appear to perform better though it is likely that a more able cohort may have driven the results. Though it is disputed that academy schools adapt their admissions policy (Thomson, 2013), this potential issue should be considered within the analysis by implementing precautionary measures. To overcome this issue, only pupils who were already enrolled in the primary school are included in the analysis; additionally, school movers are excluded from the sample. The enrolment decision is therefore exogenous to the academy conversion of the primary school.

Additionally, pupils in the treatment and control group are only compared within the same cohort; individuals in the treatment group are therefore subject to the same shocks and cohort-specific trends as individuals within the control group. Differences in pupil outcomes may therefore not be due to external factors that may affect one year group but not another.

Furthermore, a credible control group is constructed by comparing converter academies with schools that are not academies when the pupil outcomes are observed, but later convert to become converter academies. These schools may therefore have more comparable characteristics than schools that never became academies, or are not observed in the dataset to become an academy. Comparing more similar schools minimises the risk of differences in pupil outcomes being due to shocks and trends that may vary between different school characteristics.

Utilising census data from the National Pupil Database (NPD), a number of models are constructed with the analysis focusing on analysing the converter academies that converted in the 2011/12 academic year and the 2012/13 academic year whilst data from converter academies that converted in 2013/14 are utilised to construct the control group.

This chapter tests the hypothesis that converter academies improve pupil outcomes, in terms of the APS, between KS1, when children are aged 7, and KS2, when children are aged 11 and in the final year of primary schooling. While the aims of the academy programme largely related to improving educational standards within schools, the programme also allowed for schools to benefit from a greater level of autonomy, providing schools with the opportunity to innovate and target the specific needs of the school and the pupils. This hypothesis is also based upon the existing literature which suggests that whilst greater school autonomy may benefit pupil outcomes (Abdulkadiroglu et al. 2011; Böhlmark and Lindahl, 2012; Gibbons and Silva, 2011), the academy programme specifically, may show signs of assisting in pupils' educational progress (Eyles and Machin 2015; Department for Education 2014a).

The chapter provides evidence of a positive influence of converter academies upon the outcomes of primary aged pupils. The results from the difference-in-difference analysis suggest that converter academies increase the percentile rank of pupils' average point scores by between 1.3 and 2.6 percentile points, ceteris paribus. These results are based on a number of models that are adopted within the analysis; by implementing several models, the analysis considers differences in the exposure of pupils to converter academies. Relatedly, multiple cohorts of pupils are analysed as a check of the robustness of results, but also to identify whether results were equivalent across models, and therefore possibly generalizable, across cohorts and years.

An extension of the chapter is the analysis of converter academies by school neighbourhood deprivation. Though the post-2010 academy programme was less attentive to specifically failing schools within deprived neighbourhoods, since allowing for all schools to convert, the policy provides scope for the analysis of the programme based on the area of the schools. This is an interesting topic both due to the initial aims of the programme and due to the identified impact of neighbourhood deprivation upon pupil outcomes, as chapter two within this thesis identifies. The results identify large benefits of converter academies to pupils in deprived areas, though these results are confined to a small number of models. Though converter academies have a lesser impact upon pupils in the least deprived neighbourhoods, the positive effect was more consistently identified in the majority of models, which vary by cohort and year of academy conversion.

This chapter will be structured as follows; section 2 will provide a discussion of the history and background of the academies programme to date. Section 3 will review the surrounding literature. A description of the data and methodology will be provided in sections 4 and 5 respectively. The main results will be presented in section 6. Section 7 will close with a summary of the chapter's aims, findings and conclusions.

4.2 HISTORY & BACKGROUND OF THE ACADEMY SCHOOL PROGRAMME

Over the past 20 years, a number of educational systems around the world have been reformed through the introduction of new school types that deviate from the usual centrally controlled school model; instead, these schools benefit from a greater level of operational autonomy. Though varying particularly in their pre-existence, this comparable approach to the revolutionising of the education system and improving educational standards has been taken in the US, with charter schools, in Sweden with free schools or 'friskolor' and with academy schools in England.

Academy schools represent a divergence from the usual provision of state schooling in England; academies are publically funded schools that gain greater autonomy by functioning independently of local authorities. Academy schools possess greater responsibility and control over management and decisions relating to staffing, the curriculum, the length and structure of the school day and the school budget. Academy schools are predominantly conversions of pre-existing predecessor schools that convert to academy status.

The Academies Programme originated under the Labour government and was announced in 2000, with academies emerging in the English education system from 2002. The programme initially targeted three main objectives: to improve standards by raising pupil attainment, to provide inclusive, mixed-ability schools and to encourage raising standards and aspirations within the local community (National Audit Office, 2010). Initially, the programme aimed to replace underperforming secondary schools situated in socially disadvantaged urban areas with a poor GCSE attainment performance record. Formerly, academy schools were predominantly sponsored academies with a private sector sponsor. These were managed by a board of school governors, consisting of representatives appointed to the board by the sponsor and given the responsibility of school management decisions (National audit office, 2007). The conversion of schools to sponsored academies was largely used as a strategy to improve performance within these schools by converting already existing schools into academies. By 2010, 203 academies existed within the English education system (House of Commons, 2015).

Following the 2010 general election, the coalition government stated its intentions to widen the academies remit, offering primary, secondary and special schools the opportunity to seek academy status. The Academies Act 2010 consequently increased the heterogeneity of new academies by permitting all schools to apply to voluntarily convert to converter academies; this allowed for all schools to benefit from the greater autonomy of an academy without a sponsor, unlike the previously failing sponsored academies (National audit office, 2010). Under the new academies programme, schools rated outstanding by Ofsted who voluntarily applied to become academies were fast-tracked through the process of academy conversion, in an attempt to allow such schools to convert within the same year (Machin and Vernoit, 2010). Converter academies are therefore more likely to have been previously high attaining schools since initially requiring an Ofsted good or outstanding rating for academy conversion (House of Commons, 2015).

From 2010, the coalition government advised that for ‘weak’ schools, improvement could be attainable through conversion to a sponsored academy. In some cases, schools deemed ‘eligible for intervention’[19] may receive an academy order from the secretary of state, imposing steps to facilitate sponsored academy conversion. Sponsored academies therefore continue to refer to typically underperforming schools that were obliged to convert to academies. In 2014, 93% of the 1112 sponsored academies had been developed from previously failing maintained schools (House of Commons, 2015). The Academies Act therefore transformed the existing academies programme allowing for two distinct types of academy to exist within the UK schooling system.

The Academies Act 2010 also enabled new academy schools to be opened through the free schools programme. Free schools are a particular smaller category of academy that are newly opened without the conversion of a pre-existing school and are run independently of local authorities[20].

[19] A maintained school is deemed eligible for intervention under the 2006 act when “it has not complied with a warning notice and the local authority have also given the school written notice of their intention to exercise their intervention powers under… or where the school has been judged by Ofsted to require ‘significant improvement’…or ‘special measures’. Department for Education (2015)

[20] Free schools are set up and run based on proposals by groups of educators, parents, charities and others (House of Commons, 2015)

Initially, the majority of academy converters were secondary schools following the change of the academy programme; this resulted in 1,862 secondary academy converters by the end of the 2013/14 academic year. However, primary schools also began to apply for academy status, with many establishing chains of schools to convert thus allowing primary academies to benefit from pooling resources and sharing expertise (Department for Education, 2014c). By June 2015, a total of 4676 academy schools were open in England (House of Commons, 2015). The latest annual academies report indicates that by the end of the 2013/14 academic year, 2,018 primary academies existed in England; of these, 617 were sponsored academies whilst 1,401 were converter academies (Department for Education, 2015b). Previously well performing schools that voluntarily converted to become academies therefore constituted the greatest proportion of academies in the early years of the post-2010 policy (Department for Education, 2015c). Over the past five years, 80% of the growth in academies is attributed to converter academies (Eyles et al. 2015).

Following the 2016 budget, the Conservative government expressed their continued commitment to the expansion of the academy programme (Department for Education, 2016); the number of primary academy converters in England is therefore expected to continue to rise.

Greater school autonomy and school choice are argued to assist in improving educational standards and are central to the academies model. Greater school autonomy delivers greater flexibility, allowing for schools and teachers to be innovative and to provide education based on the specific needs of pupils within the school. Prospective pupils and parents may therefore be able to select schools based on taste, preference and educational needs, thus allowing for close matching of pupil needs to school provision. Furthermore, school choice is likely to influence educational standards through market mechanisms, according to standard economic theory; 'good' schools will attract a greater number of pupils and therefore funding as the school expands, while 'bad' schools that fail to attract students may be faced with closure. Market mechanisms therefore provide incentives to schools to improve and maintain good practises, especially since school performance tables that may signal school quality are made publically available; accountability is therefore a key factor in this model of schooling (Machin and Silva, 2013).

4.3 LITERATURE REVIEW

The literature may be split into two strands for discussion; firstly, there are a limited number of studies that have previously analysed the impact of academy status on pupil outcomes. Secondly, there is a field of study that focuses on analysing the impact of greater school autonomy; within this area is the analysis of US charter schools that vary from the academy programme in the UK but present similarities to the programme in the increased autonomy that is provided to schools.

The existing literature on the impact of the academy programme will be reviewed before a discussion of the remaining literature.

4.3.1 The academy programme

As the academy programme has operated since the early 2000s, with academies opening from 2002, a number of studies have previously attempted to analyse the impact of academy school status on pupil outcomes. However, as discussed within the background section of this chapter, following the election of the coalition government in 2010 and the subsequent change to the programme in the Academies Act 2010, the face of the academy school has changed dramatically. Whilst prior to 2010, academies may reliably refer to previously underperforming secondary schools that converted to sponsored academies, the alterations to policy post-2010 led to the production of both sponsored academies and converter academies, where converter academies previously attained high Ofsted ratings and typically voluntarily became academies.

Eyles et al. (2015) attempt to analyse the differences in the characteristics of Labour academies relative to the post 2010 academies using NPD data on all secondary academy conversions between the 2002/2003 and 2012/2013 academic years. The study finds that the characteristics of post-2010 academies differ significantly from Labour, pre-2010, academies; early academies are characterised by low attainment and a high proportion of disadvantaged pupils from poor family backgrounds. This is unsurprising given the remit and aims of the Labour academy policy. Coalition academies, on the other hand, have a lower share of free school meals (FSM) eligible children and experience little change in the ability composition of their intake, unlike early academies. The paper therefore warns that that generalisation and extrapolation of research findings, based on early academies, may provide a distorted and unreliable view of coalition converter academies, due to the significant difference in converting school's characteristics between the two policy periods. Though it is important to review the existing literature to evaluate the methods of analysis and data adopted, the results may not be entirely meaningful in the analysis of later converter academies.

Due to the availability of data on post-2010 academies and pupil outcomes being rather restricted, due to the short length of time since the policy change, few papers have analysed the effect of the programme since 2010. It is therefore important to take into account this change in policy when both comparing results with the existing literature and when undertaking the evaluation itself within this chapter, since it is possible that converting failing schools into academies does not have the same impact upon pupil performance as in already well performing schools. In addition, the pre-2010 programme focused on underperforming secondary schools and just as few papers have analysed the academy programme post-2010, few have analysed the programme in regards to primary school pupil performance.

Furthermore, a number of early papers analysed the impact of the academy programme at the school-level, however, as discussed by Eyles and Machin (2015), it is highly likely that the pupil composition will change following academy conversion, or more generally, following a change in institution type. The results of studies at the school-level may therefore reflect the adaptation of pupil intake following conversion.

One paper with particular relevance to this chapter is Eyles and Machin (2015) which analyses the impact of academy school conversion on pupil intake and pupil performance. Using NPD pupil level data from the 2000/01 to 2008/09 academic years, the impact of academy conversion at the secondary school level is analysed. To ensure that academy enrolment is exogenous, performance effects are examined for students who are enrolled in the secondary school before it converts to an academy. The study adopts a difference-in-differences set up and defines the control group as the students enrolled in schools that become academies after the sample period ends.

The impact of academy conversion on pupil intake is firstly considered with the quality of pupil intake being measured by the KS2 score of pupils on entry to secondary school. A differences-in-differences approach is adopted to compare the intake quality of academies before and after conversion, relative to non-academies. The results suggest that individuals with higher KS2 test scores are attracted to schools that convert to academies, thus suggesting significant differences in the pupil ability composition post academy conversion. This effect seems to grow over time but is only apparent for academies converters that were previously community schools.

The study goes on to examine the impact of academy status on pupil performance by observing KS4 (GCSE) attainment; since individuals must be already enrolled in the school when it converts to an academy, the study is limited to four years post conversion so that individuals may experience academy conversion when in their first of five years at secondary school; the set up allows for estimates to be obtained for each of the years' post-conversion separately. The years of academy school attendance therefore may vary between individuals. Not all individuals who enrolled in the school remained there to take their KS4 exams, therefore an alternative approach is taken to that analysing the impact on pupil intake. In order to estimate the Local Average Treatment Effect (LATE), selection into academy conversion, defined as treatment, is accounted for by using the intention to treat (ITT) status[21] as an instrument for the outcome which indicates whether the individual started in the school which later became an academy and if the pupil was enrolled in the school in the conversion year. The results indicate that the KS4 standardised score is significantly increased by attending an academy converter; this impact on pupil performance is increasing in the number of years of academy attendance. Relative to individuals who attend state schools that later become academies, individuals who attend academies therefore benefit from a higher value added. The paper also examines the impact of academy conversion on Ofsted inspection ratings and finds that on average, schools were likely to gain improved ratings after conversion relative to the control group used throughout the study. Overall, the study therefore suggests that academies are beneficial for educational outcomes.

Worth (2014) similarly analyses the impact of secondary school academies upon pupil performance at GCSE level; this study is one of only a few that attempts to analyse the impact of post-2010 academies. Academy schools are analysed based upon the year of conversion with the outcomes being measured in 2014, thus allowing for the duration of academy status to be examined. Data from a number of sources are adopted, including data from school performance tables, the DfE open academies list, Ofsted historical data

[21] The ITT group consists of all students who are enrolled in the school prior to conversion and are in line to complete their GCSEs at the school

and NFER's register of schools; the study utilises the data from the 2009/10 academic year until the 2013/14 academic year; all academies have been an academy school for between 2 and 4 years since the academy converters of the 2009/10, 2010/11 and 2011/12 academic years are observed. The paper uses a propensity score matching methodology to compare academies to maintained non-academy secondary schools. Schools are matched based on having similar characteristics at the time they became an academy, including the proportion of pupils gaining the gold standard GCSE outcome[22], the proportion of pupils FSM eligible and the school Ofsted rating. A number of outcome measures in are observed and are measured in 2014, including the average KS4 point score, the percentage of pupils achieving the gold standard and the average value added. A regression analysis is also undertaken to take account of differences in converter academies and the control group in terms of individual level, pupil characteristics.

The results identify an insignificant difference in the performance of converter academies relative to characteristically similar maintained schools, though in 2014, converter academies that had opened as academies for two years were found to significantly outperform schools in the control group in the attainment of the gold standard and the value added, though differences are small. The study also fails to identify a relationship between the time open as an academy and performance, when comparing earlier and later converters.

A number of largely descriptive papers, presenting non-causal impacts of academy schools have also been presented within the literature. One example is a paper by Gorard (2014) that utilises data from the NPD and Annual Schools Census (ASC) to investigate the performance and pupil intake of academies. The study echoes the findings of Eyles et al. (2015) since it suggests that sponsored academies and converter academies are vastly different with dissimilar characteristics. Additionally, the paper highlights that converter academies are much less likely to have a high proportion of FSM children whilst high percentages of converter academies are likely to be found in areas with high levels of local socio-economic segregation.

Also of particular interest to this chapter is a descriptive report from the Department for Education (2014a) which analyses the Ofsted outcomes of both primary and secondary

[22] The gold standard at GCSE refers to the attainment of 5+ GCSEs A*-C including English and maths

converter academies in 2012/13. The report indicates that relative to local authority (LA) maintained primaries, primary converter academies were more likely to maintain an outstanding Ofsted rating, while those rated good or satisfactory were more likely to gain an improved rating when inspected in 2012/13. The report also indicates that the proportion of primary converter academy pupils achieving level 4 (the expected level for KS2), or above, is much higher than in LA primary schools. Of course, considering the reviewed literature, this may reflect that higher ability pupils are attracted to converter academies since the report does not report causal effects.

4.3.2 School autonomy and school choice

Greater school autonomy is often quoted as the main benefit of academy status and one of the most common reasons for conversion amongst academy schools (Cirin, 2014); greater autonomy arguably provides academies with greater freedom in school management that aids innovation and improvement. In a survey of 2919 academies in 2013, two thirds of established converter academies reported that they believed attainment had been improved whilst academies attributed the largest improvements to be through increased collaboration and changes in school leadership.

The introduction of new types of schools into the educational system is also likely to expand school choice, which, together with autonomy, is argued to assist in increasing educational standards. With greater school choice, student preferences are likely to be more closely matched to schools, whilst market mechanisms are likely to provide incentives to schools since better schools that attract more students gain more funding (Machin and Silva, 2013).

While the research on the impact of the 2010 academy programme remains under developed, studies from other schooling systems such as in the US and Sweden, where similar school types with greater levels of managerial autonomy have been developed, provide evidence on the impact of greater school autonomy and school choice.

Since 1993, charter schools have been introduced into the American educational system; US charter schools bear a number of similarities to England’s academy schools namely due to the greater level of autonomy charter schools possess. Charter schools have control over the curriculum taught, the daily and annual school timetable and staffing decisions, just as academy schools do. Charter schools are typically newly developed schools that, whilst privately owned, remain publically funded. Since there are a number of differences between US charter schools and academies, the relevance of the charter school literature may be somewhat limited for academy schools.

A number of studies of US charter schools exploit the lottery system used to allocate students to charter schools when the school is over-subscribed. Abdulkadiroglu et al. (2011) take advantage of the random allocation of pupils to construct a quasi-experimental research design to evaluate the impact of charter school attendance on pupil attainment using data from Boston. The study also evaluates the impact of attendance at pilot schools which are similar to charter schools in terms of independence but have a lower level of autonomy. Individuals who fail to benefit from charter school attendance, due to the lottery system, are used as a control group. The results suggest that attendance at a charter school does improve test scores for middle and high school students in English language arts and in mathematics. The results are consistent with a number of other charter school studies that also utilise the lottery allocation of students in the methodology and identify positive effects of charter school attendance upon student outcomes (Hoxby and Muraka, 2009; Angrist et al. 2010).

Sweden presents an alternative education system for comparison in the analysis of England's academy schools; educational reforms in the 1990s in Sweden led to the introduction of new schools that gained greater levels of autonomy relative to existing schools, due to their independence from government control. Additionally, vouchers were introduced into the schooling system; independent schools would be funded by local government through a voucher for each pupil attracted to the school. Since free choice remained between public and independent schools, the system thereby theoretically increased the incentives for schools to raise standards and improve attainment in order to attract more students and therefore funding. Though the reform was implemented throughout the country, the number of independent schools varied across municipalities with some regions having no independent schools.

A study by Böhlmark and Lindahl (2012) uses administrative data from students born between 1972 and 1993 to identify whether changes in the average educational outcomes across municipalities were consequential of the differential growth in independent schools across municipalities. The study observes test scores at the end of compulsory education alongside high-school grades, years of schooling and university enrolment. The methodology involves estimating a basic OLS model and taking a difference-in-differences approach; changes are compared between municipalities over the time period observed to compare pre-reform years to post-reform years. The results show that even when controlling for changes in demographics, family background and characteristics of the municipality, an increased share of independent school students leads to an improvement in the educational performance. Achievement at the end of compulsory schooling, mean high-school grades and the proportion of individuals attending university are improved when the share of independent schools increases. Robustness analysis reveals that these findings are not driven by pre-reform differences in educational outcome trends between municipalities. The study therefore suggests that whilst greater autonomy may benefit schools, increased competition may also be a factor that assists in raising standards.

While attendance at an academy school is likely to influence child outcomes the experience of a change in school type while attending the school may also have an impact upon the outcomes of pupils. Within this chapter, the individuals observed in the sample experience both attendance at an academy and the conversion process. Research by Clark (2009) examines the impact of a reform in the British educational system that lead to schools becoming different types of institution. The reform allowed schools to apply to become grant-maintained (GM) high schools thus gaining autonomy by no longer being under LA control. In order for schools to gain GM status, the support of pupils and parents had to be gained by a majority vote. While this reform led to one in three high schools becoming GM schools between 1988 and 1997, a number of schools failed to gain the majority vote. The paper adopts a regression discontinuity approach by comparing performance in schools around the majority vote threshold; schools that achieved 50 percent of votes, and therefore became GM schools, are compared to schools who gained under 50 percent of votes. Using data from the annual school census 1975-2003, the study identifies that pupils attending schools that narrowly won the vote and gained greater autonomy through GM status, benefitted from significantly improved pupil performance. Though this school type is not representative of the academy school, the schools did benefit from greater autonomy. While the author argues that the results may be driven by the improvement in pupil intake quality after GM conversion, there are also suggestions that the performance gains may possibly be due to the increased resources and increased flexibility surrounding organisational changes and the use of resources.

Relatedly, Gibbons and Silva (2011) analyse the impact of attendance at primary schools that benefit from greater levels of autonomy, specifically, voluntary aided (VA) or foundation primaries. NPD data is adopted since it provides data on over one million pupils including information about each pupil's place of residence, future school attendance and academic record. The paper exploits the density of information provided by the data to estimate the impact of attendance at such schools by controlling for a multiplicity of factors that influence the propensity to attend a faith school. Individuals living in close proximity are compared to one another; the paper is also able to condition on the choice of secondary school attended to ensure characteristically similar individuals are compared within the analysis. The results from the pupil value-added model adopted indicates that there are no advantages, in terms of pupil performance, of attending more autonomous faith schools, relative to attendance at a less autonomous LEA controlled primary school. The paper suggests that pupils in faith schools are more likely to be higher in the ability distribution, signalling that more able pupils are more likely to attend more autonomous primary schools. This is a similar finding to that of Eyles and Machin (2015) and Clark (2009).

A related, earlier paper by Gibbons et al. (2008) focuses on analysing specifically the impact of school choice and competition upon pupil attainment. The authors address the possible endogeneity issue associated with school choice and pupil performance by using an instrumental variables (IV) approach. School choice is instrumented by the distance between a school and LEA boundary since it is argued that schools close to the boundaries face less competition. This is because only schools within an individual's own LEA may be attended thus pupils living near boundaries have less choice in nearby schools since other close-by schools may be outside of their LEA boundary. These pupils crowd out those who would have travelled from afar, whilst the school catchment area may be truncated near the boundary. Adopting data from the South East of England and observing the KS1 to KS2 value-added as the outcome measure of pupil performance, the results from the IV methodology fail to identify a causal relationship between school choice and competition and pupil performance. However, the results find that for primaries with autonomous governance, school competition leads to positive pupil performance gains; this effect is not identified for LEA controlled schools but for VA schools. The value-added score is increased by 1.6 points for each competitor. This effect is equivalent to 16-19 weeks progress in English and maths. Accordingly, the results suggest that increased school choice may improve standards when combined with a greater degree of school autonomy. This is found to be especially so for pupils from disadvantaged families who are eligible for FSM.

4.4 DATA

The National Pupil Database (NPD) is adopted within this chapter to identify the impact of converter academies on pupil performance. The NPD is a pupil-level database containing data on all pupils in state schools in England. The census dataset matches pupil level data to school characteristics. The NPD provides data on pupil characteristics, such as ethnicity, English as an additional language (EAL) status and free school meal status (FSM). Attainment data is also available for each pupil; attainment information is provided by the NPD from Key stage 1 to Key stage 5, allowing for pupils to be tracked over time, across schools and educational institutions.

This chapter will make use of NPD data from 2007-2014 based on the academic years observed in Table 4.1.The analysis therefore utilises the panel nature of the NPD data since the same individuals will be observed at two different points in time. Since this chapter is interested in the academy conversion of primary schools, KS1 and KS2 are specifically of interest. The data used for this analysis will include information on pupils when in the final year of KS1 aged 6 or 7 (dependent on the month of birth), and when in KS2, which marks the final year of primary schooling in the UK when pupils are aged 10 or 11.

Table 4.1 Observed academic years

KS1 ACADEMIC YEAR	KS2 ACADEMIC YEAR
2007/2008	2011/2012
2008/2009	2012/2013
2009/2010	2013/2014

This chapter additionally utilises institutional data provided by the Department for Education, which provides additional school level information on factors such as the type of institution and the school open date. The institutional data are available for all schools in the sample period and therefore provide information on changes in the school over time, such as the type of institution.

The pupil level data do provide some school characteristics but for each individual pupil, school data are only observed in the two periods in which the pupil data is reported, that is KS1 and KS2. The pupil level data are therefore combined with the institutional data, matching pupils to their schools. The matching of pupil and school data allows for school data to be observed for each individual even after they leave school and therefore after the observation period ends for the pupil. Additionally, the matching allows for school activity to be observed between KS1 and KS2. This allows for the exact academic year of changes to be identified, whereas in the pupil level data, the change would only be apparent once the pupil is observed in KS2, though a school change may have taken place in the previous academic year.

The data are therefore suitable in this pupil-level analysis of converter academies, since individuals who were attending school when the school converted to an academy may be identified. Furthermore, since the data allow for schools to be observed even after the pupil has left the school at the end of KS2, we are able to identify whether the school went on to become an academy once the pupil left the school. This information provides the basis of the methodology since it allows for individuals who experienced academy conversion to be compared to individuals from the same cohort, from similar schools who simply did not experience attendance at a converter academy as they had already left the school when the school later converted to an academy. This provides a more suitable comparison group than schools that never became academies as it is likely that schools that eventually become converter academies have more comparable characteristics. Additionally, the data are particularly valuable since they include all individuals from primary schools in England in the sample period; there is therefore no issue or questionability of the data being representative.

As indicated in Table 4.1, the NPD data utilised will span six years since observing individuals from KS1 to KS2.

The analysis will involve three cohorts; those leaving in 11/12 and 12/13 alongside those leaving in 13/14. Within these specific cohorts, a plentiful sample of individuals who attend schools that become converter academies is provided. Table 4.2 indicates the year of academy conversion experienced by individuals within the sample, given in observation terms[23]. Of course, not all pupils will be treated since some pupils will have left school by the year of conversion, for example all individuals within the sample will have completed primary school by 2015 when some primary schools convert.

Table 4.2 Year of academy conversion of school attended by sample

YEAR ACADEMY	OBSERVATIONS	PERCENT
2011	13,142	5.32
2012	54,254	21.96
2013	70,262	28.45
2014	60,028	24.30
2015	49,324	19.97
Total	247,010	

The outcome of interest within this analysis is the percentile rank of the child's average point score (APS). The APS is reported within the NPD at both KS1, when the child is aged 7, and at KS2 when aged 11 and in the final year of primary schooling. The APS in KS1 is calculated as the average score achieved in reading, writing and mathematics obtained from a teacher assessment. In KS2 the APS is based upon the average point score achieved in English, maths and science in the KS2 standardised national tests (SATs) taken at age 11 by all primary school pupils. Though these measures vary slightly in the subject that is assessed[24] since tests need to be age specific, the scores between KS1 and KS2 are comparable since providing an indication of the child's overall level of ability; it likely that there will be a high correlation between the scores achieved in individual subjects. Furthermore, the APS is used to estimate the value added score between KS1 and KS2 which is a widely adopted measure of progression between different levels of education (Gibbons et al.2013; Wilson and Piebalga, 2008). Rather than observing the APS directly, the outcome of interest within this chapter is the percentile rank of the child in terms of APS; of interest, therefore, is where the child lies in the APS distribution. The APS percentiles are estimated within both the final KS1 and the final KS2 year of the pupil; the percentile rank of the pupil is based upon their rank within their cohort and therefore indicates their percentile position amongst pupils in the

[23] This equates to 123,505 individuals whose schools become converter academies since individuals are observed in 2 periods; KS1 and KS2

[24] In KS2 English assessments, reading, writing and spelling and grammar assessed; these are therefore very similar to the assessments in KS1with the exception of science.

same school cohort who were assessed at the same time[25]. Observing the percentile rank rather than the APS score directly, allows for the three cohorts to be compared to one another; each cohort will sit a different KS2 SATs exam which may vary in difficulty; this would not be accounted for by observing the APS score. Using the percentile rank of the pupil provides an indication of where an individual sits in the ability distribution within their cohort; this rank is easily comparable across cohorts. The rank of the pupil is often observed within the education literature, for example, Machin and McNally (2008) similarly analysed the percentile reading score of pupils in the analysis of the literacy hour.

Figure 4.1 provides a density plot of the KS1 scores of both treated and control groups; the plots indicate the APS percentile distributions in the two groups before treatment. It is clear that the KS1 percentile rank distribution is similar for treated and control individuals with the exception of some points, for example at the 80th percentile, where the density of treated individuals is slightly higher than in the control group, though the distribution is very similar. The mean percentile rank among the control group is 48.39 and therefore very similar to the mean percentile rank of the treated group at 48.97.

[25] The percentile rank is based on all pupils within the cohort excluding those who are dropped from the sample due to missing data and other reasons to be specified later in this data section.

Figure 4.1 KS1 APS percentile distribution treatment and control groups

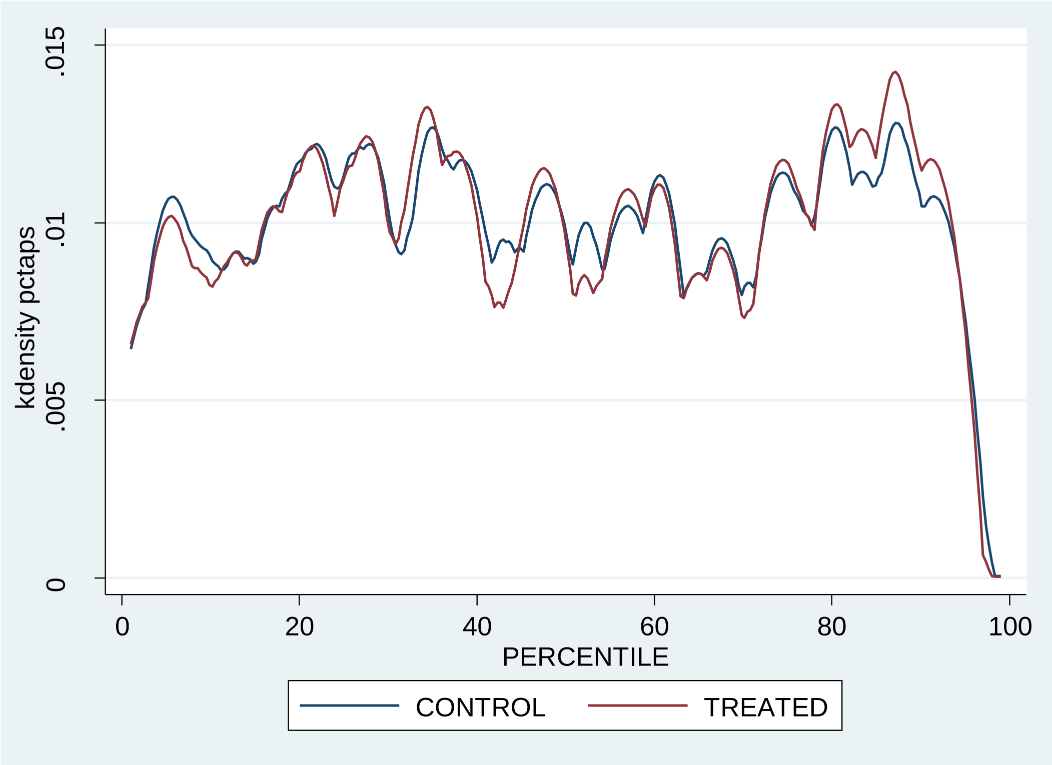
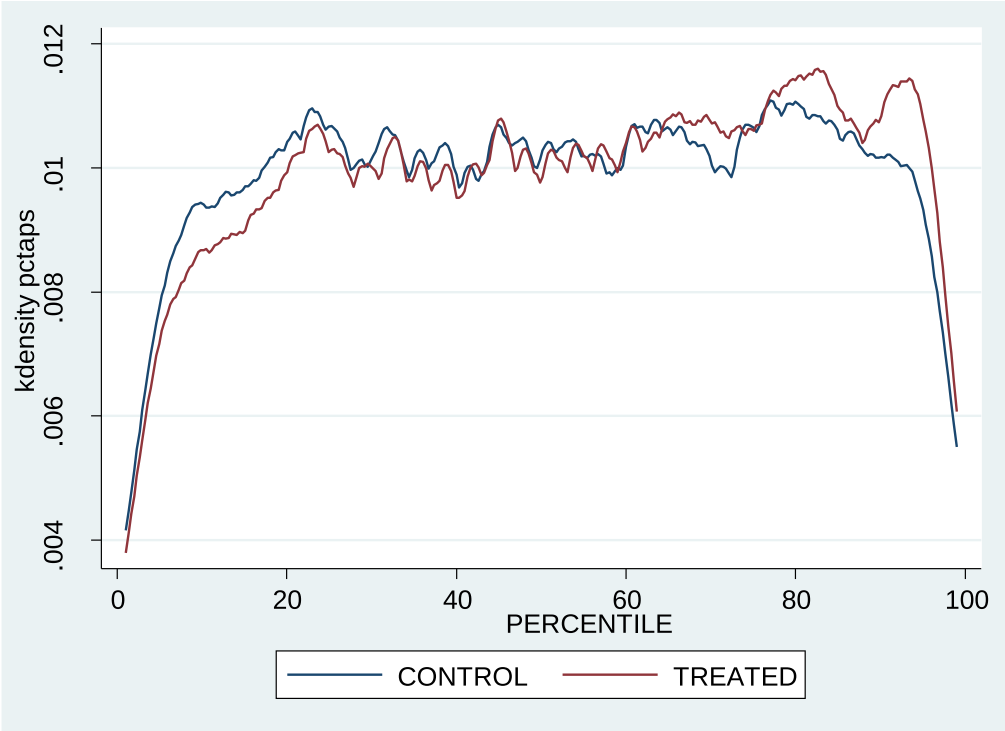


Figure 4.2 similarly provides the density plots of the APS percentile rank of individuals within the treatment group and control group, but based on the KS2 APS of individuals, therefore indicating scores after the treatment. The plot indicates differences in the distribution of the APS rank between the treatment and control group; this is especially apparent up to the 20th percentile and beyond the 60th percentile. The percentile rank density is rather similar for the treated and control groups between the 20th and 60th percentile. Beyond the 60th percentile, there is greater disparity in the distribution between treated and control groups. There is also a greater difference in the mean percentile rank of the treated and control groups, at 52.47 and 51.06 respectively. The plots therefore indicate that while individuals seem rather similar in terms of their APS rank among their cohort at KS1 before treatment takes place, once the treated group experiences academy conversion, a disparity in the distribution of the APS percentile rank begins to emerge between the treated and control groups at KS2. These raw statistics also indicate that the difference in the mean percentile score between KS1 and KS2 is greater for the treated group, with a difference of 3.5, than for the control group, with a difference in mean percentile score of 2.7.

Figure 4.2 KS2 APS percentile distribution treatment and control groups



Individuals are dropped from the sample should the unique identifier of the school attended in KS1 be different from KS2; it should be noted that the unique identifier will not change in the event of academy conversion. By dropping individuals who experience a change in the unique identifier, just over 1,100,000 observations are dropped from the sample. The unique identifier could change in three situations, thus leading to individuals being dropped from the sample. The first situation is when a school's local authority experiences a reorganisation. Secondly, the change in identifier may be due to the movement between schools between KS1 and KS2; this may simply be children moving from one primary school to another. Finally, the unique identifier will change for children who move from an infant to junior school; these children are likely to account for a large proportion of the dropped sample.

This change in the unique identifier therefore recognised individuals who should be dropped from the sample due to one of three reasons, each of which should be justified. Firstly, the individuals who experience a change in LA are dropped from the sample since a reorganisation of the LA may mean that schools are placed in different LAs or share a LA with a different set of schools following reorganisation. This could have implications for the school in terms of competition and pupil intake alongside the organisation of the school, due to changes in teachers' pay which is associated with the LA for example. Additionally, a reorganisation may prevent schools from gaining from LA initiatives in education. Generally speaking, there may be a number of impacts upon the school and therefore potentially pupil performance resulting from any changes in the structure of the LA.

Secondly, individuals who move between schools are dropped from the analysis as a school move is likely to influence the child's experience of the school and may possibly influence their outcomes. A change in scores between KS1 and KS2 for children who move schools may therefore be reflective of the school move rather than the treatment.

Finally, in dropping individuals who experience a change in the unique identifier between KS1 and KS2, children attending different infant and junior schools are removed from the sample. The reason for this is that even when the pupil attends the linked infant and junior schools, it is difficult to identify, with complete certainty, the extent to which the management, running and organisation of the school is interconnected. The difference in the organisation and management of linked infant and junior schools may vary between schools, whilst this may be a factor influencing a child's progress. In analysing the impact of converter academies, steps have been taken to ensure that the analysis carried out is based on children already attending the school. Since infant and junior schools may be seen as two separate schools, children attending these school types are excluded from the analysis to maintain consistency. The analysis of converter academies is therefore based on primary schools with combined infant and junior levels, accommodating for children aged 5-11.

Alongside the individuals that are dropped due to changes in the unique identifier, individuals are also dropped from the sample if KS1 data, including the academic year, is missing. This is because key characteristics and identifiers are unobservable. In addition, individuals who do not start KS1 and KS2 at the expected age are dropped from the sample as it is possible that a late or early start in the key stages will influence the progress and attainment of pupils. These individuals may not be comparable to all other individuals who begin the key stages at the usual age as there may be unobservable characteristics relating to their start date. Dropping these individuals leads to the loss of 523,376 observations, leaving a sample of just below one million individuals and therefore almost two million observations, since individuals are observed in both KS1 and in KS2. Table 4.3 summarises the number of observations in the sample following the loss of these individuals.

Table 4.3 Sample size summary

	SAMPLE SIZE (OBSERVATIONS)
PUPIL DATABASE:	
Original	3,624,558
Drop: Unexpected start date KS1 / KS2 & missing data.	523,376
Drop: Individuals with change in unique identifier	1,102,756
Total sample size after drops:	**1,998,426**
INSTITUTION DATABASE:	
Original	280,026
Drops	25,971
Total Sample size after drops:	**254,055**

Within the institutions data there are a number of institutions that are reported to be both academies and a different institution type within the same academic year; these observations are dropped from the analysis as it is difficult to establish from the data whether in each individual case the school misreported the institution type in one instance or whether the school did convert to an academy and then later returned to a non-academy. Dropping these observations leads to 25,971 institution observation years being removed from the sample though this eradicates the ambiguity of the type of institution in certain years. Even with the loss of these institutions, the data continue to provide over 250,000 institution observation years.

As an extension of this chapter, the analysis of academy school conversion will be undertaken when considering the area deprivation of the school, to identify whether there is a differential effect of the converter academies amongst pupils who attend schools in the most deprived areas relative to the least deprived areas. Since the data were unavailable within the NPD dataset for the years of interest, area data was sourced and matched with the Lower layer Super Output Area (LSOA) to provide the Index of Multiple Deprivation (IMD) and Income Deprivation Affecting Children Index (IDACI) associated with the school. This provides an indicator of the deprivation of the area in which the school is located and of neighbouring children. Since a higher number of primary schools exist within the UK relative to secondary schools, meaning that the average distance to school is relatively short at 1.6 miles (Department for Transport, 2014), the deprivation of the school is also likely to represent or be highly correlated with the deprivation of the child's neighbourhood. Matching in these data allows for the extension of the analysis of academy conversion but also allows for local area deprivation to be controlled for within the basic models.

4.5 METHODOLOGY

When attempting to evaluate the impact of academy conversion upon student outcomes, an issue that often arises in policy evaluation must be faced, that is, the evaluation problem. This problem arises since individuals cannot be observed in two states at the same point in time; given the states of either being treated or untreated, an individual may experience one state but not both at any time, so only one outcome per person is actually observed at any single point in time. The evaluation problem is explained thoroughly in the second chapter of this thesis.

In this chapter, the evaluation problem occurs because a treated individual's KS2 attainment cannot be observed both when the child attended an academy converter and when having attended a school that has not yet converted to an academy. For treated individuals, a counterfactual outcome is therefore required to identify the outcome where the individual does not receive treatment.

The approach taken in this chapter will involve comparing treated individuals who experience academy conversion, with individuals from the same cohort whose schools became academies but once they had left. Individuals who attend schools that do become academies but after the date which they leave provide a suitable control group, since observable and unobservable characteristics associated with academy conversion are differenced out. Schools that eventually become academies may therefore have more similar characteristics to schools that have already converted to become academies than schools that do not convert at all; this is especially so since academy conversion is not random and is a choice made by the school. This is a similar approach to Eyles and Machin (2015) who compare the outcomes of students in academy schools to the outcomes of students in state schools that convert to academies after the sample period ends.

This chapter also imitates an additional feature of the Eyles and Machin (2015) methodology by only observing individuals who already attend the primary school before academy conversion. In taking this approach, academy conversion should be exogenous to individuals who already attend the school since the school enrolment choice is made prior to conversion; individuals, or their parents, cannot therefore opt into treatment by choosing a school that is soon to become or in the process of becoming an academy. Additionally, the cohorts observed will have experienced at least three years in school before conversion, as will later be explained.

With a greater level of autonomy and independence from local authority control, schools may arguably adjust their admissions policy in order to gain a more desirable pupil intake to assist in improving academic standards (Wilson, 2011). Eyles and Machin (2015) do identify that schools which convert to academies admit higher ability students whilst Gibbons and Silva (2011) identify that more able pupils are more likely to attend more autonomous primary schools. With a higher ability intake, simply comparing all individuals who experience academy conversion to individuals in later converting schools may lead to biased results since the more able are more likely to achieve higher KS2 scores, rather than the improved attainment being due to the conversion to an academy. The approach of only identifying existing students of academy converters as treated also overcomes this potential issue since the admission of each cohort observed occurred before conversion and is therefore likely to be more similar to the intake of later converting schools in the control group.

One additional feature of the methodology adopted within this chapter is that while individuals who attend academy converters are compared to those in schools which later become academy converters, treated and control individuals within the same cohort are compared; thus, individuals in the treatment and control leave primary school in the same year. By comparing individuals within the same cohort, cohort specific effects and time effects are controlled for within the models since they are equal for treated and control individuals. This also allows for pupils to be compared within the same percentile rank since the rank relates to the APS scores of pupils in a particular cohort.

4.5.1 Difference-in-differences

A difference-in-difference methodology will be adopted in order to identify the impact of academy status on pupil performance, specifically at primary school. The approach allows for the estimation of a treatment effect, specifically, the average treatment effect for the treated. This approach is related to the before-after estimator which simply takes the difference in mean outcomes for treated individuals before the treatment occurs and after treatment; in doing so, fixed individual characteristics are controlled for in the model and the average treatment effect on the treated (ATT)[26] is obtained. This approach, however, utilises the before treatment outcome as the counterfactual for after treatment outcomes though over time there may be factors influencing the outcome; over time, it is assumed that a child's average point scores will increase as they gain more knowledge and gain learning experiences. It is therefore difficult to establish what growth in attainment should be attributed to certain factors, such as the experience of attending an academy school. One problem of the before-after estimator is therefore the assumption of temporal stability of outcomes without treatment. It should be noted that this may not be an issue within this chapter since the outcome of interest is the APS percentile rank, which may not be expected to rise over time, rather than the APS absolute score. However, this chapter will continue to implement the differences-in-differences approach as adopted within the existing literature.

The difference-in-difference estimation stems from the before-after methodology but relaxes this assumption of temporal stability, thereby taking into account possible outcome trends and changes in outcome variables over time. The difference-in-difference methodology is based on two predominant assumptions; firstly, there are common trends across the treatment and control groups, thus there are common shocks across groups and a common response to these shocks. Secondly, there is a time invariant composition assumption that suggests that participation in a programme depends on the fixed characteristics of an individual so there may be no transitory outcome shock which impacts upon participation.

Individuals who are untreated are assumed to experience the same changes over time as the treatment group so that in the absence of treatment, the treated would be subject to the same trends as the untreated. Untreated individuals therefore provide the

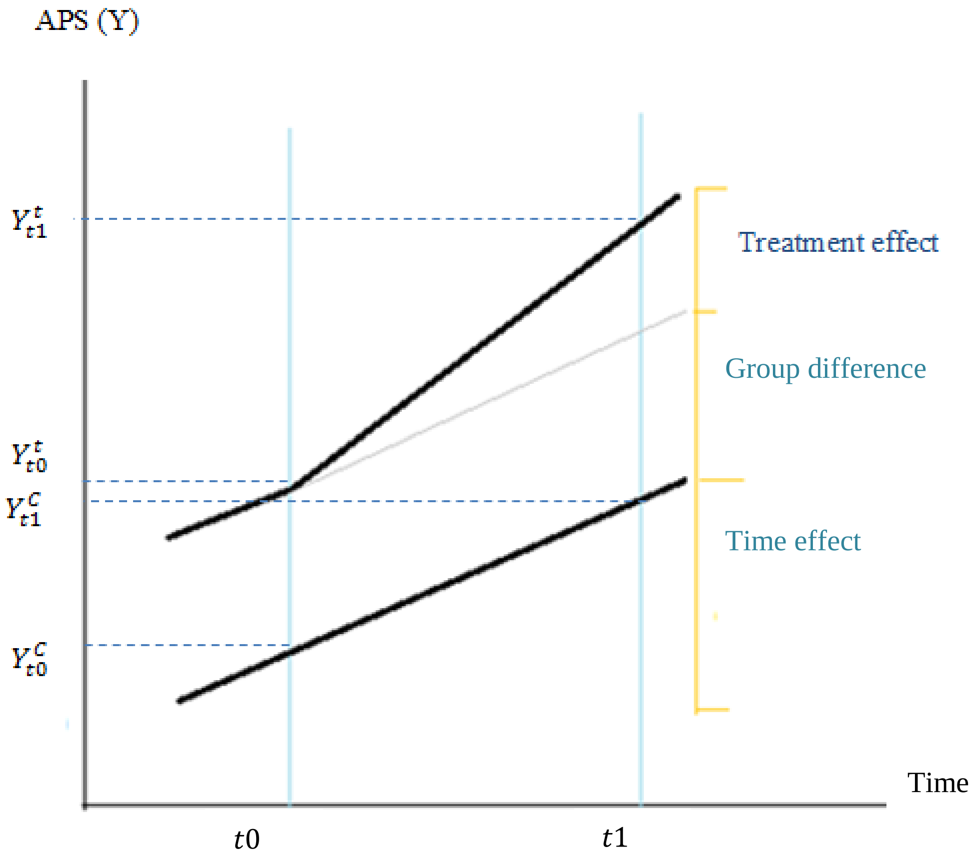
[26] Please refer to chapter 1 of this thesis for a discussion of the ATT

counterfactual outcomes of the treated, should they have not received treatment. This methodology will therefore allow for the evaluation problem to be overcome since it provides the counterfactual outcome of treated individuals in the absence of treatment.

This common trend assumption may be demonstrated diagrammatically in Figure 4.3; with the outcome, APS, on the vertical axis and time on the horizontal axis, the outcomes of the control and treatment groups are shown at Y_{t0}^{C} and Y_{t0}^{t} respectively, in the first time period *t* before treatment. After treatment, the control group is given as Y_{t1}^{C}, with the treatment outcome Y_{t1}^{t}. The grey line depicts the trend of the treatment group without the treatment that takes place after time t_0, since this follows the same upward trend as the control group. The treatment effect is therefore identified as the difference between the actual outcome (Y_{t1}^{t}) and the outcome of the treated group in the absence of treatment given by the grey line. The group difference simply indicates the difference in the treatment and control groups before treatment, which will remain following treatment. Finally, the time effect depicts the common trend over time affecting both the treatment and control groups.

This is a general depiction of the difference-in-difference model; in the analysis of this chapter, the group difference will equal zero, as will later be shown in the descriptive statistics; therefore, in the absence of treatment, not only would the trend be the same between the treatment and control group, but there should be no differences in the APS percentile rank of pupils who attend schools that become academy schools a year earlier than schools in the control group. Nevertheless, the difference-in-difference approach does account for any differences.

Figure 4.3 Difference-in-differences graphical depiction



In the simplest case with two groups and two time periods, the difference-in-difference estimator may be formally identified as:

(Eq. 4.1)

$$[E(Y_{t1}|\, D = 1) - E(Y_{t0}|\, D = 1)] - [E(Y_{t1}|\, D = 0) - E(Y_{t0}|\, D = 0)]$$

D denotes whether the individual is treated or not; Y_{t1} gives the outcome for individuals in time period 1, which is after treatment occurs, whilst Y_{t0} gives the outcome of individuals in time period 0, before treatment occurs.

The untreated group thereby controls for time effects provided that these time effects are equal for the treated and control group. Since panel data are adopted and thus the same individuals are observed over time, unobservable individuals' characteristics are controlled for within the difference-in-difference model.

4.5.2 Difference-in-difference application

This chapter focuses on the impact of attending an academy converter primary school upon child cognitive outcomes; the percentile rank of the average point score (APS) is adopted as the outcome measure of attainment. Since it is measured at two points in time while children are at primary school, both in KS1 and KS2, the percentile APS score provides a suitable outcome to adopt when using difference-in-difference analysis. KS1 APS percentile scores therefore provide the pre-treatment outcome when time is equal to 0, as in Figure 4.1 and equation1. KS2 APS percentile scores therefore provide post-treatment outcome scores when time is equal to 1.

The treatment occurs at a point in time between KS1 and KS2 when children are aged between 7 and 11 respectively. Individuals are deemed treated if they already attend a primary school which converts to an academy after the end of KS1, when the APS is measured, and before they leave the school at the end of KS2 when aged 11. The control group consists of individuals from the same cohort as the treated group whose school converted to an academy after the year they left the primary school. Since individuals in the control group are within the same cohort, they are observed in KS1 and KS2 at the exact same point in time. Unlike the usual difference-in-difference set up, there is therefore not a single point in time which marks the treatment or the introduction of a policy for all treated individuals.

The analysis will focus on comparing individuals within the three cohorts as indicated in Table 4.1; by observing three cohorts there are six main models that may be analysed; multiple models are estimated, rather than a single model, in order to determine the robustness of the results. The models to be estimated are summarised in Table 4.4. Within the first five models individuals from the treatment and control group will be from the same cohort, thus within each model the year of KS2 completion is the same for the two groups. The treatment and control group therefore vary in the date in which the school became an academy. In model A for example individuals are deemed treated if they left primary school at the end of the 11/12 academic year, and if the school became an academy in the 11/12 academic year[27]. The control group also completed KS2 at the end of the 11/12 academic year but the school that they attended did not become an academy

[27] The academic year continues to be recognised as September to August. For academy converters however, June to May is considered the same academic year since converters in July and August would not be open as academies for an entire month to students until September. E.g. an academy converting in July 2011 would be considered as converting in the 11/12 year rather than the 10/11 academic year.

until the 12/13 academic year when the control group would have already left the school. In model A, treated individuals therefore experience up to one year of attending an academy converter.

Model B is similar to model A in that the treatment group consists of individuals who leave primary school at the end of the 11/12 academic year and the school attended also converts to an academy in this year. This model varies with model A in the control group, as individuals whose school converted in 13/14 are observed, rather than 12/13 converters as in model A.

The 12AB sample simply pools model A and B so that the treatment group consists of all individuals who left primary school at the end of 11/12 and whose school became an academy in 11/12; the control group is all individuals from the same cohort whose school converted to an academy up to two years after the pupils left the primary school.

Model C varies from the other five models as treated individuals experience up to two years of attending a converter academy since in this model, treated individuals leave primary school at the end of the 12/13 academic year though the school attended may have become an academy as early as 2011.

Unlike previous models, model D examines the impact of converter academies that converted in the 12/13 academic year; previous models examined 11/12 converters.

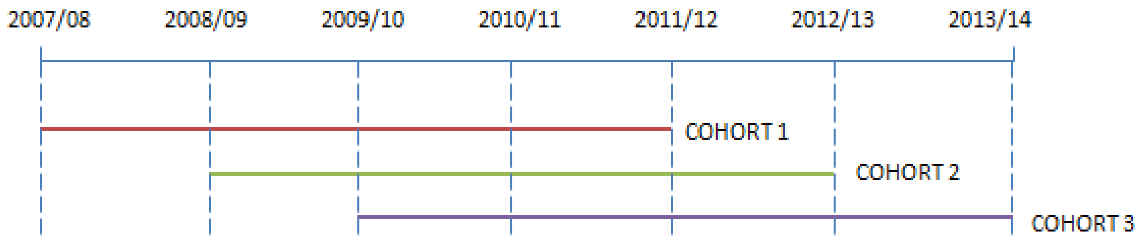
The final pooled model simply includes all treated individuals who attended school when the school converted to become an academy while the control group continues to include individuals who left school before academy conversion; these individuals may have left primary school in any year from 2012 to 2014 thus cohort dummies are entered into the pooled model to control for year and cohort specific effects. Three cohort dummies are entered for each of the year groups as indicated in Table 4.1; these dummies are based on the year that the individual began KS1 and completed KS2.

Figure 4.4 provides a diagrammatical depiction of the three cohorts that are observed within these models to aid Table 4.4. These cohorts are equally presented in Table 4.1.

Table 4.4 Models to be analysed

MODEL	BECAME ACADEMY		YEAR LEFT PRIMARY
	TREATMENT	CONTROL	
A	11/12	12/13	11/12
B	11/12	13/14	11/12
12 AB	11/12	12/13 or 13/14	11/12
C	11/12	13/14	12/13
D	12/13	13/14	12/13
POOLED	Any time while at school	After child left school	2012-2014

Figure 4.4 Cohorts observed



The first five models may therefore be represented in a general form as follows for pupil *i* in school *s* in year *t:*

(Eq.4.2)

$$PAPS_{ist} = \beta_0 + \beta_1(Treat_{is}) * (Time_t) + \beta_2(Treat_{is}) + \beta_3(Time_t) + \beta_4 X_{ist} + \beta_5 S_{st} + \varepsilon_{ist}$$

This varies from the single pooled model which may be represented as follows:

(Eq.4.3)

$$PAPS_{ist} = \beta_0 + \beta_1(Treat_{is}) * (Time_t) + \beta_2(Treat_{is}) + \beta_3(Time_t) + \beta_4 X_{ist} + \beta_5 S_{st} + \beta_6 T_t + \varepsilon_{ist}$$

PAPS gives the percentile rank of pupil i at time t in school s within their cohort according to their APS. $\beta_0$indicates the intercept term; β_1 represents the coefficient on the interaction between the treatment dummy and the time dummy and therefore indicates the treatment effect. β_2 denotes the coefficient on the treatment dummy and indicates, in the absence of treatment, the difference in the percentile rank of individuals in the treatment and control group. The treat dummy will take a value according to the model to be estimated; in model A for example, treated individuals will vary from those deemed treated in model C. The parameter β_3 is the coefficient on the time dummy equalling zero when the child is in KS1 and equalling one when the child in in KS2. The time dummy indicates how the percentile rank of pupils' changes over time, from KS1 to KS2. The parameter β_4 is a vector of coefficients on the characteristics of individual i at time t in school s affecting their distribution in the test score ranks. S_{st} is a vector of school characteristics that may influence pupils' rank in the APS distribution of their cohort. T_t, which appears in equation 3 only represents the cohort dummies which enter the model to control for cohort and year specific time effects in the pooled model. Finally, ε_{ist} denotes the error term.

It should also be noted that standard errors have been clustered at the individual level. Since the treatment variable equals 0 if the individual does not attend an academy and 1 if the individual's school converts to become an academy, the treatment variable is often serially correlated; the individual is either deemed treated or untreated throughout the data. The error term is therefore correlated over time if any unobservable characteristics determine whether or not and individual is treated, thus standard errors are likely to be biased. Clustering standard errors at the individual level is one potential solution to this issue (Cameron and Miller, 2015).

4.5.3 Characteristic controls

A number of individual and school controls are used within the model. Individual characteristics controlled for include English as an Additional Language (EAL), ethnicity, free school meal eligibility (FSM), Special Educational Needs (SEN), gender and month of birth. The school characteristics controlled for include the school institution type, the deprivation level of the school's local area and the month of academy conversion which is named 'month into'.

Firstly, a control for (EAL) enters the model as a dummy equalling one if the child is EAL and zero otherwise. Children with an alternative native language are understandably more likely to struggle in the learning environment when being taught in English, their non-native language; Sammons et al. (2007) identified that EAL is associated with more cognitive development problems for primary aged children.

Given the existing evidence on the remaining individual characteristics, it is also highly important to control for these factors. Ethnicity does seem to matter in determining a child's cognitive outcomes; Vignoles and Meschi (2010) identify that children from ethnic minority groups make greater progress than white children, while on average, females do better than males in terms of attainment. Evidence also suggests that white children in receipt of free school meals (FSM) are consistently the lowest achievers in the UK (House of Commons, 2014). Similarly, and expectedly, SEN is a consistent predictor of differences in educational attainment amongst children. Finally, the month of birth is controlled for within the model since younger children within the year group are consistently found to be academically disadvantaged, relative to children born at the beginning of the academic year (Campbell, 2013; Crawford et al. 2010); the month of birth is therefore controlled for in the model through month dummies.

Local area deprivation is controlled for within the model through dummies for the IMD decile; though the deprivation relates to the local area of the school, it is also likely to represent or be highly correlated with the deprivation of the child's neighbourhood since the average travel distance to primary school in England is 1.6 miles. The deprivation measure may also pick up the effect of the household deprivation of classmates who are equally likely to live in close distance to the school, thus, the households of these children may contribute to the IMD measure.

The type of institution is also controlled for within the model; in the second observed period, treated individuals will obviously attend an institution that is an academy converter. It is important to control for differences in institution type before conversion for all individuals and in the post-treatment period for untreated individuals since there may be differences in the characteristics of institutions that may influence attainment. For example, Silva (2006) identifies that attendance at a faith primary school leads to higher test scores in maths and English at age 11. Since faith schools have a greater level of autonomy than a number of other types of institution, the autonomy gains are likely to be much smaller when converting to an academy, relative to alternative institution types. It is important take account of differences in autonomy between institution types as greater school autonomy following conversion may be one factor that influences pupil outcomes. Three institution controls are entered into the model in accordance with the institutions observed in the data: community schools, voluntary aided schools and voluntary controlled schools with foundation primaries providing the base category[28].

Finally, the month of academy conversion will be controlled for; 'month into' indicates how many months into the academic year that the school converted, thus only schools that have converted will take a positive, non-zero value. For instance, a school converting in September will take a value of 1. Month into will equal 0 when the school has not yet converted; this will apply to both the control group observations and the treated observations in the pre-treatment period. It is important to control for the month of opening since the methodology adopted allows for schools to convert at any time within the academic year. Since conversion is likely to involve a change in leadership (Eyles and Machin, 2014), this may cause a level of disruption to the school's management and running. Additionally, since schools may convert at any time, a school converting in September will provide pupils with greater academy exposure than schools converting towards the end of the academic year; since the effect short-term exposure to converter academies are studied, it is important to distinguish and control for differential levels of converter academy exposure.

4.5.4 *Deprivation analysis*

The original Labour academy programme initially aimed to convert failing schools in deprived areas into academies; this chapter specifically evaluates the coalition policy which allows for any school to apply to become an academy. Typically, these schools were not failing schools though many may have still been located within deprived areas; this is illustrated in Figure 4.5 which will be discussed within section 6.1. It is therefore

[28] Community schools refer to state-funded primaries that are controlled by the local council. Voluntary aided schools are state-funded primary schools that receive contributions from a foundation or trust (usually a religious organisation) which has a substantial influence on the school running. Voluntary controlled primaries have less autonomy than voluntary aided schools but are similarly influenced by a foundation or trust.

interesting to identify whether there was a differential effect of the academy school conversion on schools in the most deprived areas relative to the least deprived areas given the initial aim of the original policy.

A difference-in-difference methodology will continue to be implemented when the sample is restricted to pupils who attend schools within the top 30% of deprived neighbourhoods; treated individuals will continue to be compared to untreated individuals within the same cohort who similarly attended a school within the top 30% deprived areas according to the IMD measure. This analysis will assist in identifying whether, within deprived neighbourhoods, the academy programme actually did raise academic standards. In a similar manner, the analysis will be restricted to the lowest 30% deprived neighbourhoods. The results will then be compared between the deprived and non-deprived neighbourhoods to identify whether a differential influence of academy schools exists.

4.6 RESULTS

4.6.1 Descriptive Statistics

Before presenting the main results, the descriptive statistics will be discussed, firstly, presenting the characteristics of schools that selected into becoming a converter academy, relative to those that did not become academies in the sample period[29]. Table 4.5 indicates the proportion of observations in the observed cohorts who attended a school that eventually became an academy; the data provide over 247, 000 observations[30]. As described in Table 4.4, these individuals will enter the proposed models dependent on their treatment status, cohort and the year that their school converts. Conversely, 87.6% of the sample, equal to 1,751,416 observations, attended schools that do not become converter academies within the sample period; these individuals are not included in the analysis.

Table 4.6 indicates the proportions of EAL pupils within schools that never became academies, and within converter academies. Overall, 16% of the sample is EAL; amongst the sample of individuals whose school converted to an academy, 12% were EAL whereas

[29] Between 2008-2014
[30] 247,000 observations is equal to 123,505 individuals since pupils are observed in two periods. The remainder of the results section will refer to observations rather than individuals

17% of the sample whose schools never became academies, are EAL. The statistics therefore imply that a lower proportion of EAL students experienced academy conversion.

A similar situation is portrayed in Table 4.7 which indicates the proportion of FSM students who experience converter academies relative to schools that did not become academies.

The proportion of academy converter pupils who are FSM is lower than the proportion of non-converter pupils who are FSM. The descriptive statistics therefore seem to indicate that schools that became converter academies were dissimilar to schools that did not become academies within the sample period; this is true when observing the composition of students between academies and non-academies in terms of EAL and FSM status. These descriptive statistics therefore provide support for the adopted control group within this analysis; rather than comparing converter academies to non-academies that seem to be dissimilar, converter academies are only compared with other eventual converter academies.

Turning to the sample that will be utilised within the analysis, treated individuals are defined as those who are in school when the school becomes a converter academy whilst the control group consists of those whose school became an academy once they had left. Individuals whose school did not become an academy are therefore not included in the analysis; the 247,000 individuals from Table 4.5 are therefore identified as treated or control observations within Table 4.8. There is an almost equivalent proportion of treated and control individuals contributing 55% and 45% of the sample respectively.

The IMD deciles of converters are provided in Figure 4.5. Observing the schools that become converter academies, it is clear that there is an almost equal distribution of deprivation amongst the schools that become converted[31]; from the raw data it therefore does not seem to be the case that converter academies were mostly schools that were in good, non-deprived neighbourhoods. The sample provides an almost uniform distribution of observations from each of the IMD deciles.

The institution type of the control group and the treatment group are provided in Table 4.9; these institution types obviously indicate the type of school before conversion for the treated individuals, whilst the institution type will remain the same for individuals in the

[31] The IMD decile relates to the neighbourhood of the school and is based on the 2013 IMD index

control group at the time of leaving the school; the school will later convert. Though the treatment and control group are comparable in some respects, for example community schools account for the largest proportion of institutions, the frequency and proportion of foundation schools within the treatment and control group presents a variation between the two groups; there are fewer foundation schools that remain foundation primaries than foundation schools that become converter academies. 11% of converter academies were originally foundation schools. These descriptive statistics therefore support the inclusion of institution type controls within the models; since treatment and control groups do not comprise of identical proportions of types of institutions and all institution types are pooled within the models, the impact of the type of institution on pupil outcomes should be controlled for within the model.

The remaining control variables of the models are compared between the treatment and control observations in Table 4.10. As in the previous tables, these descriptive statistics are based on all individuals who are deemed treated rather than being specific to the treatment and control groups of one specific model. These statistics are illustrative of the comparability and suitability of the control group adopted across the analysis. There are very few notable differences between the mean value of the control variables of the treated and control group. There is a slightly lower proportion of SEN pupils within the treated group and a higher percentage of white pupils in the control group while the average deprivation score (IMD decile) is higher in the treated group, signalling a higher level of deprivation; these differences are minor, however. From the statistics, it seems that there is an equal distribution of females and FSM pupils between the groups.

Table 4.11 provides the number of observations in each of the individual models of analysis. Individuals are observed at two points in time in KS1 and in KS2. Since the control group remains the same in the first three samples, the sample size is the same; this is also true for the treated group within models C and D. The pooled model contains the greatest number of observations since it includes all individuals who were treated or untreated within the 2012-2014 academic years.

Table 4.12 indicates the raw difference-in-difference statistics for each model. This is based on the APS percentile of the two individual groups. The difference indicates the change in the mean APS percentile between KS1 (before) and KS2 (after) for each group. The difference in the changes of the treatment and control group provides the raw difference-in-difference. As can be seen in the table, all difference-in-differences estimates are positive, with the exception of model D, with model C presenting the largest estimation. The estimate from model D varies from all other models since it provides a negative raw difference-in-difference estimate, suggesting that the treated group progress between KS1 and KS2 to a lesser extent than the control group whose APS percentile improves more over the observed time period.

When comparing the before estimates by model, it is clear that before treatment took place, individuals in the 12/13 cohort who attended a school that converted in 11/12 attained the highest mean KS1 percentile APS rank, this is the treated individuals from model C. The mean KS1 APS percentile is lowest among the individuals from the 11/12 cohort who attended a 12/13 converter academy; this is the control group of model A. Before treatment, when both groups were subject to the same untreated state, the treatment and control group in model A are the most dissimilar in terms of KS1 APS percentile. Of course, as previously mentioned, any group differences will be accounted for within the difference-in-difference framework.

Finally, Table 4.13 provides the raw differences in the APS scores of the three cohorts within the sample. Cohort 1 relates to the 11/12 primary leavers whilst 2 and 3 are the 12/13 and 13/14 leavers respectively. Since the APS percentile is calculated within the cohorts and is therefore only comparable within cohorts, the mean APS score is provided. The mean APS from KS1 and KS2 is presented based on all individuals within the sample from each cohort. The table simply allows for the ability or APS level of the cohorts to be compared. As the table shows, the later cohort, which includes those who complete primary school in 13/14 attain the highest APS scores both at KS1 and KS2. The APS scores increase over time by cohort, though there are not great apparent differences in the average ability level of the cohorts.

Table 4.14 similarly provides the raw differences in the APS score by cohort but also by treatment status within each cohort. Unlike Table 4.13, the mean APS scores are not generally found to increase with cohort; this is only true for the raw KS2 scores amongst control individuals who do not experience converter academies. Interestingly, KS1 scores decrease with later cohorts among treated individuals thus cohort1 obtains the highest APS score on average both at KS1 and KS2. In contrast, among the control individuals, the mean KS2 APS is lowest in cohort 1 and highest in cohort 3. Individuals from the treatment and control group are most similar in cohort 3; the mean KS1 and KS2 scores are marginally higher for the treated individuals.

4.6.2 Main results

Table 4.15 provides a summary of the main results. The results from the full model, including controls, are provided in the appendix.

4.6.2.1 Model A

Model A analyses the impact of converter academies by comparing individuals who experienced conversion in the 2011/12 academic year, while in primary school, with individuals from the same cohort whose school became a converter academy in the 2012/13 academic year, once they had left school. All individuals within this analysis completed primary school at the end of the 2011/12 academic year, thus all individuals are within the same cohort and are subject to the same time and cohort specific effects. Treated individuals receive a maximum of one year of exposure to a converter academy since the conversion may have occurred at any point in the 11/12 academic year. The results indicate that in the absence of treatment, the APS percentile rank of individuals in the treatment group is weakly significantly different to the rank of individuals within the control group at the 10% level of significance.

The time variable indicates that over time, individuals are likely to move up the percentile rank in their cohort; between KS1 and KS2 pupils move 3.7 percentiles up the APS rank within their cohort. This rank is based on the APS scores of the entire cohort of individuals included within the sample; this rank therefore includes all pupils whose school did not become an academy but excludes those who attended infant junior schools as discussed in the data section. Pupils in schools that eventually become academies move up the cohort rank which must mean that other children move down the rankings, thus, this must be the children of non-converter schools who are not included in this analysis.

The main variable of interest is the interaction of time and treatment, providing the difference-in-difference estimate of 1.9. This is a significant result indicating that converter academies have a positive impact upon the percentile test scores of children at primary school. Considering the point estimates, converter academies raised pupils' percentile APS scores by 1.9 percentile points, ceteris paribus. Relative to individuals from the same cohort whose school converted to an academy in the 12/13 academic year, treated individuals, whose schools became converter academies in the 11/12 academic year, progress in the ability ranks within their cohort to a greater extent. It is noteworthy that the difference-in-difference estimates are greater than the difference indicated in the raw descriptive statistics in Table 4.12, once controls are added. Treated individuals are exposed to converter academies for a maximum of one academic year, since the year of conversion is the year in which the pupils reach the end of KS2. If it is assumed that the benefits of conversion may increase or become more apparent with greater exposure to converter academies, as Eyles and Machin (2015) identify, then the effect may be greater for those who experience additional years of academy attendance. This will be investigated in model C.

4.6.2.2 Model B

Model B similarly analyses individuals within the same cohort who leave primary school at the end of the 11/12 academic year; the treatment group is comprised of the same individuals as model A, since these individuals experience academy conversion in the 11/12 academic year. However, the analysis varies in the control group since the school of pupils within the control group becomes an academy in 13/14 which is two academic years after the treated group.

The time between conversion of the treated and control groups' schools is greater than in model A; comparing the 11/12 to the 13/14 converters therefore continues to assist in identifying whether an impact of converter academies exists and also allows for a comparison with model A.

The results from model B signal that in the absence of treatment, the percentile rank of the treated and control group is insignificantly different. Relating back to Figure 4.1 and the surrounding discussion, it does therefore seem true that there is little group difference. Individuals are therefore comparable in terms of their APS rank since without treatment, the percentile rank of the treated group would not be significantly different from that of similar individuals in the control group. Ceteris paribus, the results indicate that the percentile rank of individuals within the analysis improves over time between KS1 and KS2 by 2.7 points; this is to a lesser extent than model A.

The main result of interest is the difference-in-difference estimate of 2.1; ceteris paribus, relative to the control group, converter academies increase treated pupils' APS percentile scores by 2.1 percentile points. Individuals who experience academy conversion between KS1 and KS2 therefore move up the ability rank of their cohort, according to APS scores, more so than individuals in the control group who do not attend a converter academy.

This is a similar yet slightly higher effect than that identified in model A. This significant result again suggests that converter academies have a positive influence upon the progress of pupils between KS1 and KS2 in comparison to pupils whose school becomes an academy but two years after their departure from the school. This effect is again the result of up to one year of exposure to a converter academy. As in model A, this result is greater than the raw difference-in-difference estimates in the descriptive statistics provided in Table 4.12.

The results so far therefore suggest that there is a greater impact of 11/12 converter academies when compared to 13/14 academies, than when compared with the 12/13 academies; this is based upon the analysis of the 11/12 cohort who have one year of exposure.

Based on conjecture alone, it could be argued that this result is identified since in the observed years, 11/12 academies were more similar to 12/13 academies that were preparing to convert; these are the academies that are observed in model A. The 12/13 converters may have already applied to become academies or may have begun to put measures in place to convert by the 11/12 academic year, when the observed pupils leave. The 12/13 converters may have not benefitted from actually becoming an academy unlike the 11/12 converters, thus possibly explaining the significant difference between the attainments of the pupils attending the 11/12 and 12/13 converters identified in model A. However, the 12/13 academies may have begun to benefit from some aspects of becoming an academy which may positively influence the attainment of their pupils, for example conversing with other converter academies to sharing expertise and pooling resources (Department for Education, 2014c). This may be less likely for 13/14 converters who form the control group of model B and did not convert for some time after the 11/12 academic year. Pupils in the control group of model B may therefore be less likely to have benefitted from the academy conversion process, unlike model A, thus producing the difference in results between the two models.

4.6.2.3 Model 12AB

In a similar manner to models A and B, model 12AB compares individuals within the 11/12 cohort; again, the schools of treated individuals became converter academies in 11/12. The control group consists of the control groups adopted in models A and B combined, thus, individuals in the control group also left school in 11/12 but their school may have become a converter academy at any time in the following two academic years from 12/13 to 13/14.

The results mirror those of models A and B; a positive and significant effect of attending a converter academy is identified upon the APS percentile rank. Considering the point estimate, the difference-in-difference estimator indicates that converter academies improve the percentile APS scores of treated pupils by 2.1, ceteris paribus. Individuals who experienced attending a school that became an academy converter are again identified as experiencing an improvement in their APS percentile rank between KS1 and KS2 to a greater extent than individuals from the control group who do not experience converter academies.

As in models A and B the results from model 12AB similarly indicate a positive and significant influence of time upon the APS percentile rank; that is, the APS percentile rank of pupils, increases between KS1 and KS2. The treatment variable indicates that the APS percentile rank of treated individuals is significantly different from the APS percentile rank of the control group in the absence of treatment, though controlled for in the model.

The results from models A, B and 12AB each present a positive and significant impact of converter academies. Each of these models was based on the same treatment group, though varying in the control group adopted for each model of analysis. The following models, C, D and the pooled model will now adopt an alternative treatment group to the previous models.

4.6.2.4 Model C

Model C defines treated individuals as those whose school became an academy converter in 11/12 in a similar manner to the previous models, though the individuals in model C will be those who leave primary school at the end of the 12/13 academic year. These individuals will therefore be from a different cohort to the previous models and will experience a converter academy for a maximum of two academic years but a minimum of one whole academic year. The control group in model C will consist of individuals from the same cohort whose school became an academy converter in 13/14, one year after they had completed KS2 and left primary school.

The results from model C are consistent with those of all previous models; the APS percentile rank of pupils in the analysis significantly improves between KS1 and KS2. Ceteris paribus, over time, individuals move up the APS percentile rank within their cohort by 2.8 percentile points. Furthermore, in the absence of treatment, treated individuals are insignificantly different to the control group, in terms of their APS percentile rank.

In model C, the difference-in-difference estimate is 1.1 which is significant at the 10% level of significant; ceteris paribus, treated individuals move 1.1 percentiles up the APS percentile rank more than in the absence of treatment. The results, again, suggest that converter academies are beneficial to pupils in terms of their APS and their ranking within the APS distribution in their cohort. Relative to the control group, treated individuals progress in the APS rank between KS1 and KS2, to a greater degree.

Model C presents one prominent difference to the previous models since treated pupils are exposed to converter academies for up to two full academic years; it is possible that the increased exposure has an additional positive influence upon pupil outcomes as the results of Eyles and Machin (2015) suggest. However, the results actually suggest that pupils benefit to a lesser extent from converter academies than the pupils in models A and B, who experienced converter academies for up to one academic year only. The result of model C is also only significant at the 10% level. This is the contradictory to expectations and the existing literature. Due to the similarities in models A, B and C, the results may not be explained by differences in the academy conversion years experienced both by the treatment and control group as differences only exist in exposure and the cohort observed. It is unlikely that the difference in results between models is due to a differential average ability of the cohorts; this is indicated within the descriptive statistics in Table 4.13, which indicate that the cohorts attain very similar APS scores both in KS1 and KS2. This is also true when observing the APS scores by cohort and treatment status, as in Table 4.14.

Arguably, the results could indicate a greater short-term beneficial impact of converter academies for pupils who are examined within the conversion year; pupils who complete primary school within the year of conversion are found to benefit from converter academies, as indicated in models A and B, to a much greater extent than pupils who are exposed to converter academies for a longer term, when exposed to academies for an additional year. This could potentially be due to the school changes experienced with conversion, providing a short-term immediate shock to student performance.

4.6.2.5 Model D

Model D is the final cohort specific analysis in which treated individuals are defined as those who experience academy conversion in 12/13 and leave in the same academic year. The control group also complete primary at the end of the 12/13 academic year but were not exposed to converter academies since their schools converted in 13/14. This model therefore varies from the previous model since the outcomes of pupils from 12/13 academies are analysed; these therefore convert one academic year later than the academies in the previous models. This cohort and control group has previously been analysed in model C. Model D varies with model C, however, since in model D pupils only experience converter academies for one year, just as in models A,B and 12AB, whereas in model C the exposure of pupils to converter academies is greater, at two years rather than one year.

The results signal a similar time effect; ceteris paribus, pupils rise 2.7 percentiles in the APS ranks of their cohort between KS1 and KS2. This is a comparable estimate to that of previous models, especially B and C which both consider the leavers from 13/14 academies as the control group.

Unlike the results from all previous models, the results of model D indicate a negative and significant coefficient on the treat variable; this suggests that in the absence of treatment, treated individuals are positioned lower in the APS percentile ranks than individuals in the control group by around 1 percentile. The APS of treated individuals is therefore lower than pupils in the control group, in the situation where no schools convert to become academies. Relatedly, the raw difference- in-difference estimates indicated by the descriptive statistics in Table 4.12, indicate a negative estimate, suggesting that the treated group progress to a lesser extent between KS1 and KS2 than the control group.

The result of the difference-in-difference estimate suggests that converter academies increase the percentile score of pupils by 2.6 more than in the absence of treatment, ceteris paribus. As in previous models, this is a positive and significant effect suggesting that individuals who attend converter academies benefit in terms of their APS. Once controls are added, this effect estimated in model D is identified as being in opposite direction to the raw difference-in-difference estimate and is greater in magnitude.

The difference-in-difference estimate in model D is a larger effect than that identified in previous models. This model is the mirror image of the 11/12 analysis in model A since the same exposure to converter academies is experienced by the treatment group and the same length of time between the control group leaving and academy conversion is analysed. There is a large difference in the difference-in-difference however, by around 1 percentile point.

Though varying in the year of academy conversion, model D bears many similarities with model C also, since adopting the same control group and analysing the same cohort; these similarities may therefore not explain the difference in results. It is therefore the treatment group that is driving this difference in results.

The results from across the models that compare individuals within specific cohorts, suggest that the 12/13 converter academies had a greater positive influence on the progression of pupils between KS1 and KS2, relative to the 11/12 academies. This could be due to these schools being somewhat 'worse' schools than the earlier converters. This is consistent with both the negative raw difference-in-differences estimate provided in Table 4.12 and the negative treat coefficient within this model; this negative and significant coefficient indicates that in the absence of treatment, treated individuals are ranked lower than untreated individuals in the control group whose schools become academies once they leave primary school. It is also consistent with the raw descriptive statistics presented in Table 4.14 which indicate that amongst treated individuals, the average APS score was lower in later cohorts. It is possible that some of the fast-tracked 'outstanding' rated early converters may have narrowly missed the 10/11 academic year and may have instead converted in the 11/12 academic years. This is less likely to be so in the 12/13 academic year since it is over a full academic year after the academy programme reform. This group of converters may therefore be less likely to consist of high proportions of 'outstanding' schools, instead, comprising more schools rated 'good' by Ofsted. It could therefore be argued that simply the gains to be made by conversion are smaller for schools that are already performing well; once treated, the 12/13 converter academies therefore provided pupils with greater levels of progression than in schools that were soon to become converter academies.

4.6.2.6 Pooled model

The final model is a pooled analysis of all individuals who experienced a converter academy between 2012 and 2014. Individuals within the control group continue to be defined as individuals whose school converted to become an academy after they had left the school; again, this may have been at any time between 2012 and 2014. Cohort dummies are entered in this analysis to control for year and cohort specific trends and effects. Relating back to Table 4.1, cohort 1 refers to the earliest cohort, who complete primary school at the end of the 11/12 academic year. Cohort 2 and 3 refers to individuals who leave primary school in 12/13 and 13/14 respectively.

The results indicate that the APS percentile rank of treated individuals is significantly different to that of the control group, in the absence of treatment. Treated individuals have a significantly higher percentile rank relative to the control group when no individuals are exposed to converter academies; this difference is controlled for. As in the previous models, the percentile rank of individuals significantly increases over time, between KS1 and KS2, as identified in all previous models.

From the pooled model results, the difference-in-difference estimate is 1.4; considering the point estimate, converter academies raised pupils’ percentile APS by 1.4, relative to individuals who are not exposed to converter academies, ceteris paribus. As in the previous models, this is a positive and significant effect but represents the smallest highly significant difference-in-difference estimate.

Alongside the main variables of the difference-in-difference analysis, it is also important to observe the other covariates within the model to see how these variables impact upon a pupil’s APS percentile rank. The full table of results is presented within the appendix of this chapter in table A4.1.

An interesting result from all models is the negative and significant impact of EAL status upon the APS percentile of pupils; EAL children are significantly lower in the ranked distribution of average point scores. This finding is in line with the existing literature which identifies that although EAL students do better in maths assessments, overall there is a negative relationship between EAL and attainment (Strand et al. 2015). Furthermore, Sammons et al. (2007) identified that EAL is associated with more cognitive development problems for primary aged children which may influence attainment.

The results from most models identify that white children do significantly worse than their non-white counterparts. Ceteris paribus, white children are estimated to be positioned between 1.5 and 2.7 percentiles below non-white children in the percentile ranks. This finding is also consistent with the existing literature; for example, Vignoles and Meschi (2010) identify that children from ethnic minority groups make greater progress than white children.

In a similar manner to EAL, FSM status negatively impacts upon attainment; considering the point estimates, being in receipt of free school meals reduces the APS percentile by between 8.1 and 8. percentile points, relative to non-eligible FSM pupils. Again, this is not a particularly unanticipated finding given the existing literature which suggests that pupils entitled to FSM begin school with lower attainment; moreover the gap in attainment persists through primary school (Strand et al. 2015).

Unsurprisingly, SEN has a negative and significant influence on the percentile rank of pupils; this is a consistent effect identified in every model. This is also a large effect signalling that SEN children are ranked around 32 percentiles lower in the APS distribution than children without SEN, ceteris paribus. The Department for Education (2014d) identifies large attainment gaps between SEN children and those without SEN in the KS1 national curriculum subjects; for example, in 2012, there was a 40 percentage point gap in reading and a 32 percentage point gap in both maths and science. This result is therefore consistent with the existing literature.

In most models, females are identified as being ranked higher than males; this is an expected result given that within the literature, females are found to attain higher test scores than males (Vignoles and Meschi, 2010). Ceteris paribus, females are ranked between 0.5 and 1.1 percentiles higher than males.

Relative to foundation primary schools, pupils attending voluntary aided schools are consistently found to rank higher in the APS percentiles. These types of institutions benefit from a greater level of school autonomy than foundation schools; voluntary aided schools in particular are found within the literature to be associated with higher pupil attainment (Gibbons et al. 2008; Gibbons and Silva, 2011). In some models, a similar effect of voluntary controlled schools is identified though this is a less consistent result.

The pupil's month of birth is also controlled for in the model since, within the educational literature, younger children within the year group are consistently found to be academically disadvantaged, relative to children born at the beginning of the academic year (Campbell, 2013; Crawford et al. 2010). The results seem to reflect this with pupils born from September to December performing significantly better than children born in January. This is a decreasing effect as the month of birth approaches January; children born in September are positioned between 3.6 and 5.4 percentiles above children born in January. December born children, on the other hand, are only placed between 1.3 and 1.8 percentiles above January born pupils, according to the models that identify a significant influence of December births. Ceteris paribus; these results are significant at the 5% level or above in only four models relative to all six models for September born children.

While the percentile rank of children born in February is predominantly insignificantly different relative to January born children, the APS rank of children born between March and August is significantly and negatively impacted. This is an increasing effect with March born children being positioned between 1.4 and 2.3 percentiles below the January born, while August born children are between 5.5 and 8 percentiles below the January born pupils.

The IMD deciles do not present such a clear picture as the month of birth effects. Only schools located in the 40% least deprived neighbourhoods according to the IMD index are consistently found to have a significant influence upon the APS rank of pupils in all models. There is a positive and significant effect of attending the schools within the least deprived neighbourhoods; this effect is increasing in the IMD decile where a higher score indicates a lower level of deprivation. There is a great difference in children in the least and most deprived neighbourhoods; children in the least deprived decile are positioned between 4.5 and 6.7 percentiles above children of the most deprived decile, ceteris paribus. Children who attend schools in the least deprived neighbourhoods therefore seem to do significantly better than children in most deprived neighbourhoods, with a negative relationship between school deprivation and APS percentile rank being apparent. These results mirror the findings of chapter 2 by similarly indicating that the children of deprived neighbourhoods perform significantly worse than the children from non-deprived neighbourhoods. The effect of school neighbourhood deprivation is less clear for those in the top 11% to 49% deprived neighbourhoods. For instance, relative to the top 10% deprived, pupils attending schools in the top 11-20% (decile 2) deprived neighbourhoods are, in some models, found to gain a worse APS rank, whilst in other models are identified to gain a significantly higher APS rank. Schools in the 21-50% deprived neighbourhoods seem to gain predominantly insignificantly different rankings to children in the most deprived neighbourhoods, with the exception of the pooled model which suggests that pupils in schools in neighbourhoods above the 10% deprivation level do significantly better than pupils in the least deprived schools.

The month into variable indicates how many months into the academic year that the treated schools converted once the school has converted. The results indicate an insignificant effect of the month of conversion upon pupil APS in all models; an increase in the month that the school converts, signalling a later conversion within the academic year, therefore insignificantly influences the pupils' APS.

Finally, the cohort dummies are entered into the pooled model only. These dummies suggest that relative to the earliest, 11/12 cohort, pupils in later cohorts are positioned at lower percentiles within the APS rank. This is a greater effect for cohort 3 as individuals are significantly ranked 1.9 percentiles below individuals of the earliest cohort, ceteris paribus.

To summarise, the main results indicate an overall positive and significant impact of converter academies, though this effect varies by model, the difference-in-difference estimates are very comparable; the estimates indicate that converter academies increase the percentile rank of pupils' average point scores by between 1.1 and 2.6, ceteris paribus.

This analysis has considered the exposure to converter academies by including models of both one and two years of academy exposure. Additionally, the analysis has considered multiple time periods of academy conversion and has undertaken the analysis utilising different samples by observing multiple cohorts.

It is important to undertake the analysis based on different cohorts and years of academy conversion since it could be possible that there is an association between school characteristics and the year of conversion; for example, early converters may be more likely to be the outstanding, fast-tracked schools. Similarly, it is vital to consider different cohorts since there may be differential impacts of converter academies on different cohorts of primary children. Despite these considerations made in the analysis, the results remain rather consistent, indicating a positive and significant impact of converter academies, thus reflecting the findings of Eyles and Machin (2015).

4.6.3 Deprivation Analysis

The sample is now split according to the deprivation level of the neighbourhood of the school attended. The analysis is carried out in a similar manner to the main models, based on only individuals who attend a school within a neighbourhood that is within the top 30%, most deprived neighbourhoods. Similarly, the sample will be restricted to only individuals who attend schools in the lowest 30% deprived neighbourhoods. This is a similar strategy to that carried out in the second chapter of this thesis which analyses the impact of neighbourhood deprivation on child outcomes.

The deprivation of the school is based on the IMD decile but as a check of robustness of these results, the sample will also be divided according the deprivation level provided by the IDACI deciles, allowing for a comparison of the results.

Table 4.16 provides the results from this analysis based on IMD scores. The results suggest that individuals from the least deprived neighbourhoods are consistently identified as benefitting from converter academies, with the exception of models C and D which identify a positive but insignificant influence. Unlike the main analysis, model A presents the greatest difference-in-difference estimate. While models A and 12AB indicate a greater benefit of converter academies within the least deprived neighbourhoods relative to the full sample, models B and the pooled model present smaller estimates. Converter academies improve the APS percentile rank of pupils in schools in the least deprived neighbourhoods by between 1.2 and 3.7 percentiles, ceteris paribus.

The results indicate that amongst the most deprived sample, there is a largely insignificant effect of converter academies on pupil performance. Only Model B indicates a significant impact of converter academies within deprived neighbourhoods. Relative to the least deprived, this estimate is rather large; ceteris paribus, converter academies significantly increase the APS percentile score of pupils by 11.8 percentile points.

The results are also presented when the sample is split according to the IDACI score as a check of robustness, as presented in Table 4.17. These results are similar to the IMD model for the least deprived neighbourhoods, since a positive and significant impact of converter academies continues to be identified in models A,B,12AB and the pooled model. These estimates are slightly smaller than the IMD estimates, with some minor differences in the significance level of these results. Interestingly, model C indicates a negative effect of converter academies though this is an insignificant effect. Within the most deprived neighbourhoods, a large, highly significant effect of primary converter academies continues to be identified in model B, but also in model 12AB, which indicates an effect of greater magnitude to the least deprived neighbourhoods but much smaller than in model B.

The results are therefore robust to a change in the deprivation measure adopted; however, the results present a rather unclear picture of the effect of converter academies upon pupils when conditioning on the deprivation level of the school. From the results, it seems that converter academies only have a positive influence on the treated from the deprived sample when the 11/12 converters are compared to the 13/14 converters, within the 11/12 cohort. Even when comparing the same year academy converters but varying the cohort observed in model C, an insignificant result is identified.

Overall, the results suggest that although in a particular year and cohort, converter academies benefitted individuals attending schools within deprived areas to a greater extent, it is the pupils of converter academies in the least deprived areas that more consistently benefit from conversion.

4.6.4 Robustness check

It is important to test the robustness of the results presented in the main analysis to identify whether the estimated impact of converter academies may be explained by alternative factors. As a check of robustness, placebo estimations from the pre-treatment period will be provided. The difference-in-difference model will be estimated as in the main analysis, however, the robustness check will be performed based upon a falsified treatment effect using a sample of individuals who left school in the pre-treatment period[32].

The robustness check will be based upon model D from the main analysis. Treated individuals attend primary schools that converted to academies in the 12/13 academic year while individuals in the control group attend 13/14 converter academies, as in model D. In the placebo model, all individuals complete KS2 in 10/11[33], and therefore leave primary school before conversion; this contrasts with model D in which individuals

[32] The observed treatment period in this chapter is 11/12 – 13/14; the pre-treatment period is therefore academic years prior to 11/12
[33] These individuals completed KS1 in 06/07

complete KS2 in 12/13. Individuals within the treatment group in the placebo model therefore receive a falsified treatment. This is illustrated in Table 4.18.

Since many of the control variables are unavailable prior to the 2006/07 academic year, the 06/07 -10/11 cohort is the earliest cohort that may be included in the placebo estimation. The placebo test could be carried out in a similar manner to models A-C; however the year of academy conversion within these models is the 11/12 academic year. It is therefore possible that the results may be susceptible to Ashenfelter's dip (Ashenfelter, 1978) which suggests that outcomes may fall directly before the treatment occurs. Model D therefore presents a more suitable model for the placebo test since in this model, treatment occurs in 12/13, two academic years after the placebo treatment. This placebo test therefore avoids any overlap between the falsified treatment and the actual treatment.

The placebo test is therefore a test of whether the impact of converter academies identified may be due to a pre-existing difference between the treatment and control group, prior to treatment. Specifically, the test will identify whether there were pre-existing differences in the outcomes of pupils who attended 12/13 converter academies, relative to pupils who attended schools that converted in 13/14, that may be perceived as the treatment effect within the main analysis. If the results are robust, there should be an insignificant difference-in-difference estimate in the placebo model; this result would imply that there were no pre-existing differences in the outcomes of the treatment and control group prior to the actual year of treatment, while pupils could not be impacted by academy conversion, since leaving primary school (in 10/11) before their school converted.

The results presented in Table 4.19 indicate an insignificant difference-in-difference estimate suggesting that the falsified treatment has an insignificant influence on the outcomes of the treatment group relative to the control group. The results therefore pass the placebo test, signifying that there are no significant pre-existing differences in the outcomes of the pupils who attend schools that become academies in the year of the treatment group and those who attend schools that convert in the control group year, that could have alternatively explained the difference in outcomes that was attributed to academy converters in the post-treatment period.

The insignificant treat coefficient indicates that in the absence of treatment there is an insignificant difference between the treatment and control group. The results also show a significant time effect, as in the main models, showing that individuals move up the percentile rank over time.

The results from the placebo test therefore imply that the results from the main analysis are robust, since when a falsified treatment is entered into the model in place of an actual treatment, the results fail to identify a significant influence upon outcomes.

4.7 CONCLUSIONS

This chapter has examined the impact of converter academies upon pupil outcomes at the primary school level. Specifically, the chapter has attempted to identify how being exposed to a converter academy impacts upon pupils' percentile rank within their cohort, according to their average point scores. Individuals who attended primary schools that converted to academies were deemed as treated individuals whilst pupils who completed KS2 before their primary school converted to became an academy were used within the control group. These individuals provided a suitable control group, since the schools attended were likely to be more similar since eventually all becoming converter academies, compared with individuals in school that did not become converter academies within the sample period.

Data from the National Pupil Database provides data on all pupils that are in school in England within the years examined in the analysis. The data therefore provided a suitable source of data on students required for this analysis since they provide large samples of individuals that experienced academy conversion, or attended a primary school that became an academy converter after KS2. The data utilised within the analysis was from 2008-2014 which covered three main cohorts of pupils who completed primary school in 2012, 2013 or 2014.

A difference-in-difference methodology was adopted in line with existing papers within the surrounding relevant literature (Eyles and Machin, 2015; Wilson, 2011; Böhlmark and Lindahl 2012). This methodology assisted in overcoming the evaluation problem, allowing for the outcomes of the control group to be compared to those of the treatment group, thus allowing for a treatment effect to be identified.

The main analysis involved a number of models which varied by the cohort observed or the year of academy conversion of the treatment or the control group. The models predominantly examine the impact of converter academies that converted in the 11/12 academic year though one model analyses the impact of 12/13 converters whilst a pooled model analyses converters at any time between 2012 and 2014.

By only observing individuals who had already enrolled in the primary school, the approach attempted to overcome the potential problem of endogeneity by ensuring the enrolment decision was exogenous to the academy conversion of the school.

The results of the main analysis indicate a positive and significant impact of converter academies upon the percentile rank of primary pupils; the estimates indicate that converter academies increase the percentile rank of pupils' average point scores by between 1.4 and 2.6 percentile points, ceteris paribus. This positive impact varies between models but is a consistent finding throughout the analysis suggesting that regardless of the year of conversion between 2012 and 2014, there is a positive and beneficial effect at the primary level. This finding mirrors that of Eyles and Machin (2015) who analyse secondary converter academies pre-2010. There are few papers that focus on the academy programme to compare the results with. However, the results are in line with the existing research on school autonomy that suggests that pupil outcomes are positively influenced by a greater level of school autonomy (Gibbons et al. 2008; Abdulkadiroglu et al. 2011; Clark, 2009).

To my knowledge, this chapter contains the first non-descriptive results to be presented indicating the impact of coalition primary converter academies on pupil outcomes.

A robustness check is carried out in the form of a placebo test by analysing the effect of converter academies on a sample of individuals who attended schools that eventually became academies within the sample period, but once this sample had completed primary school. Running model D with a sample of individuals from the 10/11 cohort, the placebo test results found an insignificant impact of converter academies, or the placebo effect, upon the APS percentile rank of pupils. As expected, the results therefore indicated signifying that there were no significant, pre-existing differences in the outcomes of the treatment and control group that could have alternatively explained the difference in outcomes that was attributed to academy converters in the post-treatment period.

The analysis also estimates the impact on converter academies by neighbourhood deprivation to identify whether the pupils attending schools within the least deprived neighbourhoods benefit differentially to those in the most deprived areas. The results consistently suggest a positive impact of converter academies the least deprived neighbourhood sample in most samples. For the most deprived sample, a positive and significant effect of converter academies is identified consistently in only one model, when altering the deprivation measure, though a very large effect is identified. Schools in deprived neighbourhoods therefore do not seem to benefit from academy conversion in all periods. It is only when the 11/12 academy converters are compared to the 13/14 academy converters and the 11/12 cohort is observed, that a significant influence of converter academies is identified for the deprived neighbourhoods.

From this research, the results overall suggest a positive role of the academies programme in improving the progress of primary school children. From this analysis, the change in policy made by the coalition government, allowing for all schools to become academies, therefore seems a positive transformation which has, at least in the case of the 11/12 and 12/13 converter academies, begun to assist in improving the future outcomes of pupils, by advancing their progress between KS1 and KS2.

The results suggest the 12/13 converter academies had a greater positive influence than the 11/12 converters. It is possible that this is due to the later converting schools being 'worse' schools as later converters were less likely to be the fast tracked 'outstanding' schools. The results and descriptive statistics do provide evidence of this. It should be noted, however, that at all points in time, converters must meet certain requirements including a 'good' or 'outstanding' rating, so even if the later converters contained a lower proportion of 'outstanding' ratings, the converters were still 'good' schools.

Though it is possible that the exposure to converter academies will influence the impact upon pupil outcomes, it is vital to consider how children who experience the conversion process and experience a short amount of time in an academy converter will be influenced. This is especially of concern given that only a small proportion of primary schools have converted at present; thus, millions of pupils may experience the process of academy conversion within their primary schools within coming years.

Future research should make use of additional years and cohorts of primary pupils; since at the time of analysis, the latest pupil outcomes available related to 2014, there was little scope to analyse the most recent converter academies. This is because, in this methodology, the control group consists of individuals who attend schools that convert at least a year after the treated group, thus academy conversion is observed in the latest years of data whilst pupil outcomes are observed in the previous year. Additional years of data will therefore allow for further years of academies to be analysed.

Furthermore, though converter academies represent the greatest proportion of academy schools in England, sponsored academies continue to exist and continue to be opened. It is therefore equally as interesting to examine the impact of these academies. However, due to the pronounced expansion of primary converter academies within England and the plans for all schools to convert, it is imperative to identify the impact of this particular policy primarily.

Table 4.5 Observations attending eventual converter academies within the sample period 2007-2014

SCHOOL BECAME AN ACADEMY	FREQUENCY	PERCENT
0	1,751,416	87.64
1	247,010	12.36
Total	1,998,426	

Table 4.6 Proportion of EAL observations in eventual converter academies and non-converter academies

SCHOOL BECAME AN ACADEMY	EAL	
	0	1
0	1,457,106 (83.35%)	291,134 (16.65%)
1	218,296 (88.45%)	28,506 (11.55%)

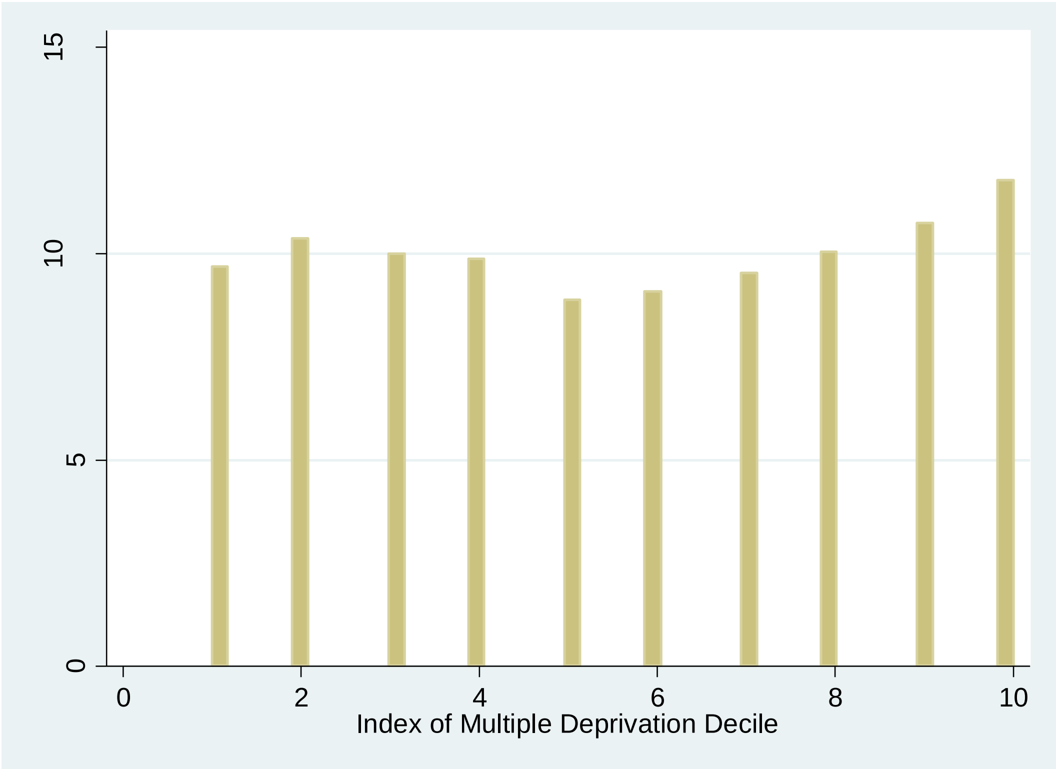
Table 4.7 Proportion of FSM observations in eventual converter academies and non-converter academies

SCHOOL BECAME AN ACADEMY	FSM	
	0	1
0	1,258,532 (71.96%)	490,482 (28.04%)
1	187,784 (76.04%)	59,168 (23.96%)

Table 4.8 Treated and untreated individuals within the sample

TREATED	FREQUENCY	PERCENT
0	111,506	45.14
1	135,504	54.86
Total	247,010	

Figure 4.5 IMD decile of eventual academy converters



Note: Higher values denote a lower level of deprivation

Table 4.9 Institution type

INSTITUTION TYPE	CONTROL	TREATED
Community school	63,580 (57.02%)	83,108 (61.33%)
Voluntary aided school	33,146 (29.73 %)	27,452 (20.26%)
Voluntary controlled school	10, 376 (9.31%)	10,028 (7.4%)
Foundation school	4,404 (3.95%)	14,916 (11.01%)
Total	111,506	135,504

Table 4.10 Mean of remaining control variables

VARIABLE	TREATED	CONTROL
EAL	0.12	0.11
White	0.82	0.84
FSM	0.24	0.24
SEN	0.20	0.21
Female	0.49	0.49
Month of birth	6.5	6.6
IMD decile	5.7	5.5
Month into	3.2	0

Table 4.11 Sample size of each model of analysis

MODEL:	CONTROL (0)	TREATED (1)
A	22,767	17,482
B	19,307	17,482
12AB	42,075	17,482
C	19,688	17,417
D	19,688	22,845
Pooled	110,783	134,278

Table 4.12 Raw before, after and difference-in-difference estimates by model

	BEFORE	AFTER	DIFFERENCE	DID
Model A control	47.62	51.29	3.67	
Model A treated	50.55	54.72	4.17	0.5
Model B control	48.46	51.13	2.67	
Model B treated	50.55	54.72	4.17	1.5
Model 12AB control	48	51.22	3.22	
Model 12AB treated	50.55	54.72	4.17	0.95
Model C control	48.51	51.26	2.75	
Model C treated	51.11	55.15	4.04	1.29
Model D control	48.51	51.26	2.75	
Model D treated	47.43	48.51	1.08	-1.67
Pooled model control	48.39	51.08	2.69	
Pooled model treated	48.97	52.47	3.5	0.81

Table 4.13 Mean pupil KS1 and KS2 scores by cohort

COHORT	KS1APS	KS2APS
1	15.17	28.56
2	15.26	28.64
3	15.3	28.96

Note: Cohort 1 refers to individuals leaving primary school in the 11/12 academic year, cohort 2 leave in 12/13 and cohort 3 in 13/14.

Table 4.14 Mean pupil KS1 and KS2 scores by cohort and treatment status

COHORT		KS1APS	KS2APS
1	Control	15.43	29.03
	Treated	15.70	29.53
2	Control	15.56	29.08
	Treated	15.62	29.36
3	Control	15.52	29.38
	Treated	15.55	29.48

Table 4.15 Summary of main results from difference-in-difference analysis for all models

	(1) MODEL A	(2) MODEL B	(3) MODEL 12AB	(4) MODEL C	(5) MODEL D	(6) POOLED MODEL
Time * treat	1.887*** (0.669)	2.085*** (0.741)	2.073*** (0.638)	1.077* (0.563)	2.604*** (0.812)	1.431*** (0.251)
Time	3.735*** (0.167)	2.715*** (0.184)	3.265*** (0.124)	2.753*** (0.182)	2.744*** (0.182)	2.733*** (0.077)
Treat	0.616* (0.329)	0.419 (0.348)	0.583** (0.293)	0.375 (0.347)	-1.023*** (0.314)	0.894*** (0.145)
N	40,249	36,789	59,557	37,105	42,533	24,5061

Note: Controls also include: EAL, ethnicity, FSM, SEN, gender, school type, month of birth, school neighbourhood IMD decile, school open month. Pooled model also includes cohort controls. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.16 Summary of results from difference-in-difference analysis by school neighbourhood deprivation according to IMD.

TREAT:	30% LEAST DEPRIVED	N	30% MOST DEPRIVED	N
A	3.747*** (1.306)	13,658	1.454 (1.700)	10,703
B	2.022* (1.051)	14,067	11.844*** (2.954)	9,583
12AB	2.761*** (0.971)	19,985	2.605 (1.641)	17,636
C	0.546 (0.923)	14,108	1.424 (1.418)	9,782
D	2.008 (1.393)	12,429	-2.136 (1.515)	15,079
Pooled	1.163*** (0.437)	79,843	-0.312 (0.490)	73,673

Note: The time*treat coefficient is provided only
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.17 Summary of results from difference-in-difference analysis by school neighbourhood deprivation according to IDACI.

TREAT:	30% LEAST DEPRIVED		30% MOST DEPRIVED	N
A	2.563*** (1.075)	12,693	2.005 (1.635)	10,454
B	1.915** (1.092)	12,788	9.742*** (2.309)	9,118
12AB	1.841* (1.015)	18,508	3.232** (1.570)	16,434
C	-0.173 (0.958)	12,936	0.248 (1.258)	9,379
D	1.948 (0.444)	11,741	0.417 (1.681)	13,417
Pooled	0.900** (0.444)	72,859	0.216 (0.504)	70,512

The time*treat coefficient is provided only
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.18 Placebo test sample summary

MODEL	BECAME ACADEMY		YEAR LEFT PRIMARY
	TREATMENT	CONTROL	
D	12/13	13/14	12/13
PLACEBO	12/13	13/14	10/11

Table 4.19 Placebo test results summary

VARIABLES	PLACEBO
Time * treat	0.366 (0.140)
Time	2.769*** (0.184)
Treat	0.234 (0.309)
N	41,232

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

4.8 APPENDIX

Table A4.1 Full model results with all controls

	(1) Model A	(2) Model B	(3) Model 12AB	(4) Model C	(5) Model D	(6) Pooled model
Time * treat	**1.887*****	**2.085*****	**2.073*****	**1.077***	**2.604*****	**1.431*****
	(0.669)	**(0.741)**	**(0.638)**	**(0.563)**	**(0.812)**	**(0.251)**
Time	**3.735*****	**2.715*****	**3.265*****	**2.753*****	**2.744*****	**2.733*****
	(0.167)	**(0.184)**	**(0.124)**	**(0.182)**	**(0.182)**	**(0.077)**
Treat	**0.616***	**0.419**	**0.583****	**0.375**	**-1.023*****	**0.894*****
	(0.329)	**(0.348)**	**(0.293)**	**(0.347)**	**(0.314)**	**(0.145)**
EAL	-3.073***	-2.190***	-2.650***	-2.936***	-3.557***	-2.667***
	(0.746)	(0.848)	(0.632)	(0.792)	(0.705)	(0.305)
White	-0.387	-1.561**	-0.749	-2.713***	-2.089***	-1.581***
	(0.627)	(0.708)	(0.535)	(0.671)	(0.615)	(0.258)
FSM	-8.726***	-8.742***	-8.664***	-8.127***	-8.660***	-8.513***
	(0.377)	(0.402)	(0.306)	(0.395)	(0.351)	(0.152)
SEN	-32.649***	-32.256***	-32.305***	-32.779***	-31.902***	-31.996***
	(0.347)	(0.367)	(0.285)	(0.378)	(0.345)	(0.145)
Female	0.186	0.758**	0.505**	1.092***	0.982***	0.903***
	(0.302)	(0.316)	(0.248)	(0.318)	(0.297)	(0.124)
Community	0.679	0.089	0.611	0.611	1.185	0.020
	(0.537)	(0.638)	(0.484)	(0.453)	(0.730)	(0.214)
Voluntary aided	2.689***	1.688**	2.075***	1.525***	1.638**	2.208***
	(0.607)	(0.699)	(0.532)	(0.535)	(0.773)	(0.240)
Voluntary controlled	2.863***	0.267	1.985***	0.195	2.132**	0.481
	(0.752)	(0.848)	(0.641)	(0.724)	(0.885)	(0.303)
MOB Feb	-0.605	-0.865	-0.612	-1.321*	-1.062	-1.235***
	(0.753)	(0.786)	(0.618)	(0.787)	(0.737)	(0.311)
MOB March	-0.501	-2.321***	-1.409**	-2.963***	-2.074***	-1.872***
	(0.727)	(0.767)	(0.604)	(0.769)	(0.725)	(0.303)
MOB April	-2.507***	-2.393***	-2.663***	-3.144***	-3.095***	-2.896***
	(0.733)	(0.773)	(0.602)	(0.778)	(0.728)	(0.304)
MOB May	-2.632***	-2.574***	-2.882***	-3.684***	-3.785***	-3.801***
	(0.727)	(0.754)	(0.592)	(0.766)	(0.710)	(0.299)
MOB June	-4.192***	-4.407***	-4.058***	-5.524***	-4.847***	-4.919***
	(0.725)	(0.769)	(0.598)	(0.770)	(0.718)	(0.301)
MOB July	-4.767***	-5.213***	-4.857***	-5.232***	-5.655***	-5.449***
	(0.737)	(0.769)	(0.605)	(0.768)	(0.720)	(0.298)
MOB August	-5.539***	-5.846***	-5.710***	-7.239***	-8.079***	-6.681***
	(0.722)	(0.750)	(0.591)	(0.755)	(0.713)	(0.297)
MOB September	5.463***	4.506***	5.148***	3.716***	3.567***	4.770***
	(0.723)	(0.753)	(0.595)	(0.761)	(0.716)	(0.300)
MOB October	3.552***	3.260***	3.326***	3.176***	2.744***	3.181***
	(0.740)	(0.775)	(0.608)	(0.767)	(0.726)	(0.302)
MOB November	2.242***	2.129***	2.296***	0.399	1.719**	2.304***
	(0.728)	(0.774)	(0.602)	(0.772)	(0.719)	(0.307)
MOB December	1.534**	1.551**	1.787***	0.808	1.283*	1.432***
	(0.736)	(0.776)	(0.606)	(0.792)	(0.734)	(0.307)
IMD 2	1.595**	-0.294	0.906	-2.209***	-0.391	1.402***
	(0.726)	(0.746)	(0.553)	(0.729)	(0.596)	(0.271)
IMD 3	-0.928	-0.679	-0.552	-0.600	0.579	1.377***
	(0.728)	(0.746)	(0.552)	(0.748)	(0.592)	(0.278)
IMD 4	0.888	1.769**	0.773	0.441	-1.687***	0.607**
	(0.737)	(0.754)	(0.569)	(0.738)	(0.632)	(0.278)
IMD 5	0.963	-0.029	0.625	-0.709	0.801	1.950***
	(0.721)	(0.821)	(0.577)	(0.810)	(0.665)	(0.289)
IMD 6	2.532***	3.822***	2.829***	1.221	1.607**	3.006***
	(0.749)	(0.766)	(0.580)	(0.743)	(0.664)	(0.291)

IMD 7	2.181***	2.287***	1.815***	2.102***	1.896***	3.387***
	(0.772)	(0.775)	(0.594)	(0.775)	(0.671)	(0.289)
IMD 8	3.991***	4.054***	4.185***	3.421***	3.886***	4.673***
	(0.788)	(0.731)	(0.577)	(0.731)	(0.644)	(0.285)
IMD 9	5.375***	3.445***	4.379***	3.655***	3.126***	5.230***
	(0.741)	(0.741)	(0.574)	(0.724)	(0.677)	(0.283)
IMD 10	6.135***	5.750***	6.043***	4.522***	5.771***	6.769***
	(0.715)	(0.737)	(0.557)	(0.726)	(0.641)	(0.278)
Month into	-0.040	-0.028	-0.032	0.094	-0.023	-0.027
	(0.062)	(0.062)	(0.062)	(0.062)	(0.059)	(0.024)
Cohort 2						-0.759***
						(0.157)
Cohort 3						-1.907***
						(0.167)
Constant	54.946***	57.159***	55.619***	59.264***	57.781***	56.504***
	(1.110)	(1.205)	(0.919)	(1.087)	(1.147)	(0.449)
Observations	40249	36789	59557	37105	42533	245061
Adjusted R^2	0.320	0.312	0.315	0.301	0.300	0.299

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

CHAPTER 5 : CONCLUSIONS

This thesis has analysed the impact of three factors, each relating to a pupil's environment outside of their household, upon pupil outcomes. The impact of neighbourhood deprivation, class setting by ability and primary converter academies have been examined, in turn, when observing pupils in compulsory schooling within the UK. Throughout the three chapters that form this thesis, both cognitive and non-cognitive measures of pupil outcomes have been observed in an attempt to contribute to the existing literature and the gaps that are present within it. Additionally, the thesis adds to the existing research on the determinants of pupil outcomes by addressing factors that are external to the household and the family and are instead related to a pupil's external environment. Specifically, the thesis focuses on the impact of a pupil's neighbourhood and factors associated with a pupil's school life, such as school-level and national-level policies. At present the evidence on these external influences is less clear cut and arguably underdeveloped within the economics of education field.

5.1 THESIS SUMMARY

The first empirical chapter, chapter 2, examined the impact of neighbourhood deprivation upon pupils' cognitive outcomes, measured by the probability of obtaining the benchmark GCSE outcomes; 5 GCSEs A* to C and 5 GCSEs A* to C including English and maths. Though a number of existing studies have explored neighbourhood deprivation, few have adopted econometric techniques to overcome the methodological issues that face researchers, namely, the evaluation problem, selection bias and the establishment of causality. One method that assists in overcoming these issues is propensity score matching, though this method is not extensively employed in existing studies of neighbourhood effects. Of the few papers within the economics literature, only a limited number of studies have evaluated neighbourhood effects in a static manner, that being, without evaluating individuals who move between neighbourhoods. Additionally, few studies analyse the impact of neighbourhood deprivation as defined by deprivation indices.

Using data from the initial three waves of the Longitudinal Study of Young People in England, relating to 2003 to 2006, the chapter contributes to the existing literature by employing a propensity score matching methodology to analyse the impact of neighbourhood deprivation, defined by the IDACI score, upon the outcomes of individuals who live in a deprived neighbourhood for at least three years. The chapter firstly examines the impact of neighbourhood deprivation for all individuals within the sample before exploring the differential impact of neighbourhood deprivation according to parental education.

The initial results reveal an overall negative impact of living in deprived neighbourhoods upon the probability of obtaining the benchmark GCSE outcomes; individuals living in deprived neighbourhoods are around 4 percentage points less likely to obtain five GCSEs A* to C, relative to individuals living in non-deprived neighbourhoods and are around 6 percentage points less likely to obtain five GCSEs A* to C including English and maths.

When evaluating the impact of neighbourhood deprivation according to parental education, where educated parents are defined as those educated to at least post-16 level, the results similarly signal a negative impact of neighbourhood deprivation for both individuals with educated and uneducated parents. The results interestingly reveal that the estimated neighbourhood effect is significantly larger for individuals with educated parents than those with uneducated parents. The results therefore signal that the penalty associated with neighbourhood deprivation upon the educational attainment of residents, is greater for individuals with educated parents who would benefit to a greater extent by living in a non-deprived neighbourhood, relative to individuals of uneducated parents.

This is an interesting result that presents a novel finding within the neighbourhood effects literature, identified when adopting an alternative methodology and deprivation measure to the existing studies.

Remaining within the field of education and similarly focusing on the determinants of pupil outcomes, Chapter 3 explores the impact of the school-level policy of class setting. At present, this policy is determined at the school level though it has received attention from the Conservative government, with sources suggesting that class-setting may be made compulsory within all schools (BBC, 2014).

This chapter focuses on class setting in mathematics in the primary school specifically since, within primary schools, maths is the subject for which most pupils are exposed to the implementation of setting. As indicated within the existing literature that considers the psychological impacts of class setting upon children, class setting may be detrimental to the attitudes and behaviours of children (Eilam and Finegold, 1992; Kutnick et al., 2005). This is particularly important since anti-school attitudes developed at a young age, for example at primary school, may determine the child's performance and attitudes throughout their academic years.

This chapter therefore considers how setting impacts upon the behaviour of primary school children, using the responses of both parents and teachers in the SDQ to measure behaviours. Specifically, using data from the Millennium Cohort Study, the chapter firstly considers how being set for maths, relative to not being set, influences behaviour. A fixed effect methodology is adopted for this analysis to overcome the issue of unobserved heterogeneity. The analysis also examines how the level of the maths set in which the child is placed influences behaviour; an alternative methodology is adopted in this part of the analysis since there is a possible issue of endogeneity; an instrumental variables approach is therefore adopted.

Results from the fixed effects analysis suggested that the act of setting children for maths in primary school was beneficial for girls' behaviour. Specifically, the teacher reported internalising behaviour of girls was improved when setting in maths was implemented. Contrastingly, for boys, the analysis indicated that setting was detrimental to the behaviour of boys.

The results from the analysis of the level of the set indicates that for girls, teacher reported internalising behaviours were improved by being placed in the lowest set for maths, whilst for boys, being placed in the lowest set for maths was found to increase the internalising behaviour score reported by the parent.

The findings therefore signal some significant impacts of setting in primary school which contrasts with a number of papers within the relevant literature that suggest little influence of setting (Whitburn 2001; Barker Lunn 1970; Kulik and Kulik 1992; Ireson and Hallam, 2005). However, given the variability of the significance of results, the findings do not provide a clear representation of the overall impact of class setting upon behaviour.

The results interestingly suggest a non-uniform impact of class setting by gender; these findings therefore suggest that simply evaluating policies that may influence child behaviour as a pooled sample may not provide a clear picture. The behaviour of children should be examined by gender since it is clear that whilst there are differences in the types of behaviour that the genders exhibit (McNeish and Scott, 2014; Leadbeateret al. 1999), there are also differences in the reaction of behaviour to policies or alternative shocks.

The chapter contributes to the existing literature since class setting is a policy that is generally unaddressed within the economics literature; very few existing studies therefore adopt econometric techniques to overcome the surrounding methodological issues. The results of the analysis suggest that class setting may be a determinant of pupil non-cognitive outcomes, thus, when examining pupil outcomes, it may be important to identify and control for class setting practises that pupils may be exposed to. This chapter provides a foundation for future work to further analyse class setting policies or similar ability grouping practises.

The final empirical chapter, chapter 4, examines the effect of primary converter academies upon the progress of pupils between Key stage 1 and Key stage 2. Following the election of the coalition government in 2010, the pre-existing academies policy was transformed, allowing for all schools to apply to voluntarily convert to an academy converter, when meeting specific requirements. Though a number of studies have examined the impact of academies, few have done so after 2010 when within the English education system, the number of converter academies rapidly expanded. Furthermore, the House of Commons (2015) reported a need for research into the impact of primary academies since there are very few studies that analyse the academies programme at the primary level.

This chapter therefore contributes to the literature in two ways; firstly, by providing an analysis of post-2010 academies, and secondly, by focusing on primary schools within this analysis.

Following a similar strategy to Eyles and Machin (2015), the chapter utilises data from the National Pupil Database and adopts a difference-in-difference analysis to evaluate the impact of converter academies upon children who were already enrolled in school before the conversion; this allows for the enrolment choice to be exogenous to the conversion. The approach taken within this analysis also involves defining a credible control group by comparing pupils who experienced converter academies with pupils whose schools become academies but after they leave the school. Furthermore, in order to control for cohort specific effects, individuals who are treated are compared with control individuals from the same cohort.

Within the difference-in-difference analysis, a number of models are defined that allow different years of converter academies to be compared, using various cohorts of pupils. With the NPD data adopted providing data on KS2 pupil outcomes from 2012-2014, the analyses estimate the impact of converter academies that converted in the 11/12 and 12/13 academic years. 13/14 converters are utilised in the control group.

The results suggest that converter academies have a positive and significant impact upon pupil progress; this is true in all models examined. Pupils therefore seem to benefit from attending a converter academy when compared with individuals who attend schools that are soon to convert to academies.

The analysis also considers whether there is a differential impact of converter academies by the level of neighbourhood deprivation in the area in which the school is located. The deprivation of the school's neighbourhood is also likely to reflect the pupil's neighbourhood deprivation due to the, on average, short travelling distance to primary schools in England. This is of interest given that in chapter 2, neighbourhood deprivation was found to negatively influence pupil outcomes. This finding therefore motivates the analysis of chapter 4 which attempts to identify whether converter academies have influenced pupil outcomes to a greater extent in deprived neighbourhoods than non-deprived neighbourhoods. This is also interesting given that the initial aims of the academies programme, prior to the 2010 transformation, was to target schools within deprived areas. Since it was proposed that such schools would benefit from the greater levels of autonomy provided by academy status, it would be expected that, despite the change in policy, schools within deprived neighbourhoods continue to benefit from autonomy.

The results indicate that in some models, pupils attending schools that are located in deprived neighbourhoods have significantly greater progress than pupils in schools that do not convert; the effects were greater than those identified for converter academies within non-deprived neighbourhoods. However, whereas in deprived neighbourhoods only a small number of models signalled a significant impact of converter academies, in non-deprived neighbourhoods, pupils attending converter academies out-performed the pupils of non-converters in most models. The results from the non-deprived sample therefore revealed more consistent effects of converter academies.

5.2 POLICY IMPLICATIONS AND FUTURE RESEARCH

The results of chapter 2 indicate that cognitive pupil outcomes are negatively impacted by residing in a deprived neighbourhood; the neighbourhood in which an individual resides, and the characteristics of that neighbourhood, should therefore be considered and controlled for within analyses of pupil outcomes. Whilst the results indicate that all children are negatively influenced by neighbourhood deprivation, it is children of educated parents who lose out to a greater extent by residing in a deprived area. The findings therefore suggest that targeting children based upon their socio-economic status, may fail to aid those who suffer from the effects of neighbourhood deprivation to the greatest extent. Efforts should be made to ensure that whilst children are targeted based upon residing in a deprived area or coming from families that do not have a history of continuing in education after the compulsory leaving age, that the children of educated parents are equally the focus of the many policies that aim to improve educational attainment, if the government does intend to provide the opportunity for all children to reach their full potential (Department for Education, 2015c).

The analysis of the impact of class setting does reveal, in some models, a significant impact of the school-level policy upon child behaviour, though some results are inconsistent across specifications. Whilst indicating a mix of both positive and negative effects, the results do not provide clear policy implications. Due to the results, it is suggested that more research should be undertaken in order to establish the relationship between class setting and child behaviour since it is important to understand whether this school level policy benefits or hinders the non-cognitive development of children. It is also imperative that research is undertaken to continue to fill the gap in the economics literature which at present provides little input in the ability grouping debate. Specifically, future work could be focused on the impact of class setting in alternative subjects; due to small sample sizes, this analysis was not carried out within chapter 3. Additionally, it may be interesting to observe how alternative pupils' outcomes are influenced by class setting, for example test scores at age 11.

The findings of the final empirical chapter present a promising illustration of primary converter academies. In the main analysis, all results identified a positive and significant impact of converter academies upon pupil outcomes. Of course, at an early stage in the research process and with only a few years of the transformed academy programme having being experienced, there is certainly scope for future work. Firstly, additional years of data should be exploited; in the methodology adopted in chapter 4, additional data would provide both an extra year of converter academies to evaluate and would also allow for further control groups to be utilised. This is the main concern of future research since it is important to understand whether results are generalizable across the years of conversion. Secondly, there is scope to expand the analysis by including infant and junior schools if data are available on the interrelatedness of junior schools and feeder infant schools. Finally, and more generally, future research should also attempt to analyse the impact of post 2010 secondary converter academies; though fewer in number than primary converter academies, secondary academies represent a greater proportion of secondary schools in England.

To summarise, each chapter within this thesis individually contributes to the literature on neighbourhood effects, ability grouping and academy schools within the UK; at present, given the lack of consensus within the literature on each of these topics in the economics of education field, there is certainly scope to build upon the existing research. In its entirety, the thesis sheds light on the determinants of pupil outcomes that relate to the wider environment in which a pupil comes into contact, outside of their household. The analyses presented within this thesis therefore provide an initial step towards closing the gap in these areas of the economics of education literature.

BIBLIOGRAPHY

Abadzi, H. (1984) Ability Grouping Effects on Academic Achievement and Self-Esteem in a Southwestern School District. *The Journal of Educational Research* 77(5), 287-292

Abdulkadiroglu, A. Angrist, J., Dynarski, S., Kane, T. & Pathak, P. (2011) Accountability and flexibility in public schools: evidence from Boston's charter and pilots. *Quarterly Journal of Economics* 126, 699-748

Ainsworth, J. (2002) Why does it take a village? The mediation of neighbourhood effects on educational achievement. *Social Forces* 81(1), 117-152

Alexander, N., & Campbell, E. (1964). Peer influence on adolescent educational aspirations and attainments. *American Sociological Review* 29(4), 568-575

Anders, J. (2012) *Using the Longitudinal Study of Young People in England for researching into Higher Education access.* Department of Quantitative Social Science, working paper no.12-13

Angrist, J., Dynarksi, S., Pathak, P. & Walters, C. (2010) Inputs and impacts in charter schools. KIPP Lynn. *American Economic Review* 100, 239-43

Ashenfelter, O. (1978) Estimating the effect of training programms on earnings. *Review of Economics and Statistics* 60: 47-57

Aslund, O., Edin, P., Fredriksson, P. & Grönqvist,. H. (2011) Peers, Neighborhoods, and Immigrant Student Achievement: Evidence from a Placement Policy. *American Economic Journal: Applied Economics* 3(2): 67-95.

Atkinson, A. Burgess, S. Gregg, P. Propper, C. & proud, S. (2008)*The impact of classroom peer groups on pupil GCSE results*. Centre for Market and Public Organisation working paper no. 08/187

Barker, J. Smith, F. Morrow, V., Weller, S. Hey, V. & Harwin, J. (2003) *The impact of out of school care: A qualitative study examining the views of children, families and playworkers.* Department for Education and Skills, Research Report RR446

Barker-Lunn, J. (1970) *Streaming in the primary school: a Longitudinal Study of Children in Streamed and Non-streamed Junior Schools*. Slough, NFER

BBC (2014) *Should England make setting and streaming compulsory*? Retrieved 19th October 2016 from http://www.bbc.co.uk/news/education-29051923

Becker, G. (1964) Human capital: a theoretical analysis with special reference to education. Columbia University Press, New York.

Betts, J. & Shkolnik, J. (2000) The effects of ability grouping on student achievement and resource allocation in secondary schools. *Economics of Education review* 19 (1), 1-15.

Biggs-Gowan, M. Carter, A. & Schwab-Stone, M. (1996) Discrepancies among mother, child, and teacher reports: examining the contributions of maternal depression and anxiety. *Journal of Abnormal Child Psychology* 24(6), 749-65

Blatchford, P. Bassett,P., Brown, P., Martin, C.,Russell, A. & Webster, R. (2009) *Deployment and impact of support staff.* Institute of education, University of London.

Boaler, J. (1997) Setting, social class and survival of the quickest. *British Educational Research Journal* 23, 575-595

Böhlmark, A. & Lindahl, M. (2012) *Independent Schools and Long-run Educational Outcomes: Evidence from Sweden's Large-scale Voucher Reform*. IZA discussion paper 6683

Bohnert, K. & Breslau, N. (2008) Stability of psychiatric outcomes of low birth weight: A longitudinal investigation. *Archives of General Psychiatry* 65(9),1080-1086.

Bolster, A., Burgess, S., Johnston, R., Jobes, K. Propper, C & Sarker, R. (2007) Neighbourhoods, households and income dynamics: a semi-parametric investigation of neighbourhood effects. *Journal of Economic Geography* 7, 1-38

Bradley, S. & Taylor, J. (2008) *Do peers matter? Estimation of peer effects from pupil mobility between schools.* Department of Economics. Lancaster.

Brattbak, I. and Wessel, T. (2012) Long-term Neighbourhood Effects on Education, Income and Employment among Adolescents in Oslo. *Urban Studies* 1, 1-16

Brooks-Gunn,J., Duncan, G., & Lawrence, J.(1997) Neighborhood Poverty: Context and Consequences for Children. Volume 1 York: Russell Sage Foundation Press

Bryson, A., Dorsett, R. & Purdon, S. (2002) *The use of propensity score matching in the evaluation of active labour market policies.* Department for Work and Pensions working paper no.4

Caliendo, M. & Kopeinig, S. (2005) *Some practical guidance for the implementation of propensity score matching.* IZA discussion paper 1588

Cameron, C. & Miller, D. (2015) A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources* 50 (2), 317-372

Campbell, T. (2013) *In-school ability grouping and the month of birth effect: preliminary evidence from the Millennium Cohort Study*. Centre for Longitudinal Studies (CLS) working paper 2013/1

Campbell, T. (2013) *In-school ability grouping and the month of birth effect: preliminary evidence from the Millennium Cohort Study*. Centre for Longitudinal Studies (CLS) working paper 2013/1

Canadian Paediatric Society (2004). Maternal depression and child development. *Paediatric Child Health* 9 (8), 575-583

Carrell, S. & Hoekstra (2010) Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone’s Kids. *American Economic Journal: Applied Economics* 2(1), 211–228

Centre for Longitudinal Studies, Institute of Education (2014) *Millennium Cohort Study (MCS) webinar.* 25th November 2014

Chanfreau, J. (2015) *Out of school activities in the primary years: analysis of Millennium Cohort Study* (Powerpoint presentation) Presented at Centre for Longitudinal Studies Cohort Studies Research Conference 2015 17th March 2015.

Cheshire, P. (2007) *Segregated Neighbourhoods and Mixed Communities: A critical analysis.* Joseph Rowntree Foundation. York.

Chetty, R., Hendren, N. & Katz, L. (2016) The effects of exposure to better neighbourhoods on children: new evidence from the Moving to Opportunity Experiment. *American Economic Review* 106(4): 855-902

Chowdry, H. Crawford, C. & Goodman, A. (2010) *The roles of attitudes and behaviours in explaining socio-economic differences in attainment at age 16.* Institute for Fiscal Studies working paper 10/15

Christofides, L. Hoy, M., Milla, J. & Stengos, T. (2012) *Grades, aspirations and post-secondary education outcomes.* IZA discussion paper 6867

Cirin, R. (2014) *Do academies make use of their autonomy?* DfE research report 366

Clark, D. (2009) The performance and competitive effects of school autonomy. *Journal of Political Economy* 117, 745-783

Collins, C. & Gan, L. (2013) *Does sorting students improve test scores? An analysis of class composition.* National Bureau of Economic Research working paper 18848.

Connelly, R. (2013) *Millennium Cohort Study Data Note: Interpreting test scores.* Centre for Longitudinal Studies, Institute of Education

Coren, E., Hutchfield, J., Iredale,W. & Thomae, M.(2010) Child Outcomes-Focused Evaluation Final Data Collection Report. Canterbury Christ church University.

Cowen, G. & Wilton, M. (2008). *Impacts and Outcomes of the Neighbourhood Renewal Fund. Communities and local government.* Retrieved 1st June 2014 from http://www.yorkconsulting.co.uk/uploads/78_Impacts_and_Outcomes_Neighbourhood_Renewal_Fund.pdf

Crawford, C. Dearden, L. & Meghir, C. (2010) *When you are born matters: the impact of date of birth on educational outcomes in England*. Institute for fiscal studies working paper w10/06

Cutler, D. & Glaeser, E. (1997) Are ghettos good or bad? *The Quarterly Journal of Economics* 112(3), 827-72

Dearden, L., McGranahan, L. & Sianesi, B. (2004) *The role of credit constraints in educational choices: evidence from NCDS and BCS70*. CEE discussion paper no. 48

Deng, Suo, Huang, J. Minchao, J. & Sherraden, M. (2013) Household assets, school enrolment, and parental aspirations for children's education in rural China: Does gender matter? *International Journal of Social Welfare* 23(2), 185-194

Department for Education (1993) *Improving primary education – patten*. DfE news 16/93 London. Department of Education

Department for Education (2011) *Academies Annual Report. Academic Year 2011/2012.* Department of Education

Department for Education (2012). *Attainment at Key Stage 4 by Pupils in Academies 2011*. DfE Research Report 223

Department for Education (2014) *GCSE and equivalent attainment by pupil characteristics in England, 2012/13.* Retrieved 7th May 2014 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/280689/sfr05_2014_text_final.pdf

Department for Education (2014a) Performance of converter academies: an analysis of inspection outcomes 2012 to 2013. Research report Department for Education

Department for Education (2014b) *Academies annual report. Academic year: 2012 to 2013.* Department for Education

Department for Education (2014c) *Phonics screening check and national curriculum assessments at key stage 1 in England, 2014.* Department for Education

Department for Education (2015a) *Schools causing concern. Statutory guidance for local authorities.* Department for Education

Department for Education (2015b) *Reading: the next steps. Supporting higher standards in Schools.* Department for Education

Department for Education (2015c) *Academies annual report. Academic year: 2013 to 2014*. Department for Education

Department for Education (2016) Next steps to spread educational excellence everywhere announced. [Press release]. [Accessed 25 May 2016]. Available from https://www.gov.uk/government/news/next-steps-to-spread-educational-excellence-everywhere-announced

Department for Government and local communities (2008) *The English Indices of Deprivation 2007.* Retrieved 6th May 2014 from http://webarchive.nationalarchives.gov.uk/20120919132719/http://www.communities.gov.uk/documents/communities/pdf/733520.pdf. Department for Government and local communities

Department for Transport (2014) *National Travel Survey 2014: Travel to school.* Department for Transport

Dex, S.& Joshi, H. (2004) *Millennium Cohort Study First Survey: A User's Guide to Initial Findings*. Retrieved 18th March 2015 from http://eprints.ioe.ac.uk/5933/1/MCS1_A_Users_Guide _To_Initial_ Findings.pdf

Dickerson, A. & Mcintosh, S. (2013) The impact of distance to the nearest education institution on the post-compulsory education participation decision. *Urban Studies* 50(4), 742-758

Dubrow, N., Edwards, S. & Ippolito, M. (1997) Life stressors, neighbourhood disadvantage and resources: A focus on inner-city children's adjustment. *Journal of Clinical Child Psychology* 26, 130-144

Dunne, M. Humphreys, S. & Sebba, J. (2007) *Effective Teaching and Learning for Pupils in Low Attaining Groups*. Department for Children, Schools and Families report DCSF_RR011

Eilam, B. & Finegold, M. (1992) The heterogeneous class: a solution or just another problem? *Studies in Educational evaluation* 18(2), 165-78

Ellen, I. & Turner, M. (1997) Does neighbourhood matter? Assessing recent evidence. *Hosing Policy Debate* 8(4), 833-866

Ermisch, J. & Pronzato, C (2010) *Causal Effects of Parents' Education on Children's Education*. Institute for social and economic research 2010-16

Eyles, A. & Machin, S. (2015) *The introduction of Academy schools to England's education*. IZA discussion paper no.9276

Eyles, A., Machin, S. & Silva, O. (2015) *Academies 2: The new Batch.* CEP mimeo

Fauth, R. Parsons, S.& Platt, L. (2014) *Convergence or divergence? A longitudinal analysis of behaviour problems among disabled and non-disabled children aged 3 to 7 in England.* Institute of Education working paper no. 14-13

Finn, J., Pannozzo, G., & Achiles, M. (2003) The 'why's' of class size: student behaviour in small classes. *Review of Educational Research,* 73(3), 321-368

Gallop, K., Rose, n. Wallace, E., Williams, R., Cleary, A., Thompson, A., Burston, K., Frere-smith, T. Dangerfield, P., & Tietz, S. (2013) *Millennium Cohort Study fifth sweep (mcs5).* Centre for longitudinal studies.

Gamoran, A. (2002) *Standards, inequality & ability grouping in schools.* Centre for Educational Sociology Briefing 25.

Gennetian. L., Sciandra, M., Sanbonmatsu, L., Ludwig, J., Katz. LF., Duncan, GJ., Kling, JR. & Kessler, RC (2012). The Long-Term Effects of Moving to Opportunity on Youth Outcomes. *Cityscape* 14(2), 137-68.

Gibbons, S. & Silva, O. (2011) Faith primary schools. Better schools or better pupils? *Journal of Labor Economics* 29(3), 589-635

Gibbons, S. & Telhaj, S. (2012) *Peer effects: Evidence from Secondary School Transition in England*. IZA Discussion Papers 6455.

Gibbons, S. (2002) *Neighbourhood effects on educational achievement: evidence from the census and national child development study*. Discussion Paper LSE18, Centre for the Economics of Education, London School of Economics

Gibbons, S. Silva, O. & Weinhardt, F. (2014) *I (don't) like the way you move: the disruptive effect of residential turnover on student attainment*. Mimeo LSE

Gibbons, S., Machin, S. & Silva, O. (2008) Choice, competition and pupil achievement. *Journal of the Europena Economic Association*. 6(4), 912-947

Gibbons, S., Machin, S., & Silva, O. (2012) *Valuing school quality using boundary discontinuities*. Centre for Economics of Education

Gibbons, S., Silva, O. Weinhardt, F. (2012) Everybody needs good neighbours? Evidence from students' outcomes in England. *The Economic Journal* 123(571), 831-874

Goodman R (1997) The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry* 38, 581-586.

Goodman, S., Rouse, M., Connell, A., Broth, M., Hall, C. & Heyward, D. (2011) Maternal Depression and Child Psychopathology: A Meta-Analytic Review. *Clinical Child and Family Psychology* review 14(1), 1-27

Gorard, S. (2014) The link between academies in England, pupil outcomes and local patterns of socio-economic segregation between schools. *Research in Papers Education* 29 (3), 268-84

Gould, E., Lavy, V. & Paserman, D. (2011) Sixty years after the magic carpet ride: the long run effect of early childhood environment of social and economic outcomes. *Review of Economic Studies* 78(3), 938-73

Goux, D. & Maurin, E. (2006) Close neighbours matter: Neighbourhood effects on Early performance at school. *Economic Journal, Royal Economic Society* 117(523), 1193-1215.

Gregg, P., Propper, C & Washbrook, E. (2008) *Understanding the Relationship between Parental Income and Multiple Child Outcomes: a decomposition analysis.* CMPO working paper 08/193

Grilli, L.& Rampichini, C. (2011) *Propensity scores for the estimation of average treatment effects in observational studies*. Training sessions on causal inference Bristol June 2011

Gutman, L. & Akerman, R. (2008) *Determinants of aspirations. Centre for Research on the Wider Benefits of Learning.* Research report 27

Guttmanova, K. Szanyi, J & Cali, P. (2008) Internalizing and Externalizing Behavior Problem Scores: Cross-Ethnic and Longitudinal Measurement Invariance of the Behavior Problem Index. *Educational and Psychological Measurement.* 68(4), 676-694

Hadow report (1931) *The primary school* (London, HMSO)

Hagan, J. & Foster, H. (2001) Youth violence and the end of adolescents. *American Sociological Review* 66 (6), 874-899

Hallam, S. & Parsons. (2014) The impact of streaming on attainment at age seven: evidence from the Millennium Cohort Study. *Oxford Review of Education,* 40(5), 567-589

Hallam, S.Ireson, J. Lister, V. & Chaudhury (2003) *Ability grouping practices in the primary school: A survey*. Institute of Education

Hanushek, E. (2000). *Evidence, Politics, and the Class Size Debate*. Mimeo., Hoover Institute, August

Hanushek, E., Rivkin, S. Kain, J. (2005) Teachers, schools and academic achievement. *Econometrica* 73: 417-458

Harding , D. Gennetian, L. Winship, C. Sanbonmatsu, L. & Kling, J. (2010) *Unpacking neighbourhood influences on educational outcomes: setting the stage for future research.* NBER Working Paper No. 16055.

Harding, D. (2003) Counterfactual models of neighbourhood effects: the effect of neighbourhood poverty on dropping out and teenage pregnancy. *American Journal of Sociology* 109(3), 676-719

Heckman, J. (2000) Policies to foster human capital. *Research in Economics* 54(1), 3-56

Heckman, J., Ichimura, H. & Todd, P. (1998) Matching as an econometric evaluation estimator. *Review of Economic Studies* 65, 261-294

Hersch, J. (1996) Smoking, seat belts, and other risky consumer decisions: differences by gender and race. *Managerial and Decision Economics* 17(5), 471–481.

Hill , R., Griffiths, W. & Lim, G (2007) *Principles of econometrics.* Wiley.

Hirsch (2007) *Experiences of poverty and educational disadvantage.* Joesph Roundtree Foundation, York.

HM Treasury (2016) *Budget 2016*. London: House of Commons

Holland, P. (1986) Statistics and causal inference. *Journal of the American statistical association* 81 (396), 945-960

Hope, S. Pearce, A. Whitehead, M. & Law, C. (2014) Family employment and child socioemotional behaviour: longitudinal findings from the UK Millennium Cohort Study. *Journal of Epidemiology Community Health* 68, 950-957

House of Commons (2015) *Academies and free schools: Fourth Report of Session 2014–15.* London: House of Commons

House of Commons (2011) *Behaviour and Discipline in Schools*. London: House of Commons HC 516-II

House of Commons (2014*) Underachievement in Education by White Working Class Children. First Report of Session 2014-15*. London: House of Commons

House of Commons (2005) *Higher standards, better schools for all*. London: HMSO

Hoxby, S., & Muraka, S. (2009) *Charter schools in New York City: who enrols and how they affect student achievement.* National Bureau of Economic research working paper 14852

Hutchings, M., Francis, B. & De Vries., R. (2014). *Chain Effects: the Impact of Academy Chains on Low Income Students.* Retrieved 4th February 2016 from http://www.suttontrust.com/wp-content/uploads/2014/08/chain-effects-july-14-final-1.pdf

Individual Data. Cited in Brooks-Gunn,J., Duncan, G., & Lawrence, J.(1997) *Neighborhood Poverty: Context and Consequences for Children. (*Volume 1) York: Russell Sage Foundation Press

Ireson, J. & Hallam, S. (2005) Pupils' liking for school: Ability grouping, self-concept and perceptions of teaching. *British Journal of Educational Psychology* 75, 297–311

Ireson, J. & Hallam, S. (2009) Academic self-concepts in adolescence: relations with achievement and ability grouping in schools. *Learning and instruction* 19, 201-213

Ireson, J. Clark, H. & Hallam, S. (2001) *Ability grouping in Education*. London, sage.

Ireson, J. Hallam, S. & Hurley, C. (2005) What are the effects of ability grouping on GCSE attainment? *British Educational Research Journal* 31(4), 443-458.

Ireson, J. Mortimore, P., & Hallam, S. (1999a). The common strands of pedagogy and their implications in Ireson, J. & Hallam, S. Ability grouping in Education (pp 30-39, pp46-51).

Ireson, J., Hallam, S. Mortimore, P., Hack, S. Clark, H. & Plewis, I. (1999b) *Ability grouping in the secondary school: the effects on academic achievement and pupils' self-esteem*. Retrieved 24th March 2015 from http://www.leeds.ac.uk/educol/documents/00001359.htm

Jackson, B. (1964) *Streaming: an education system in miniature*. London: Routledge and Kegan Paul.

Jianakoplos, N. & Bernasek, A. (1998) Are women more risk averse? *Economic Inquiry* 36(4), 620–630

Katz, L., Kling, J. & Liebman, J. (2001) Moving To Opportunity In Boston: Early Results Of A Randomized Mobility Experiment. *Quarterly Journal of Economics,* 116(2), 607-654.

Kelly, Y., Kelly, J. & Sacker, A. (2013) Changes in Bedtime Schedules and Behavioral Difficulties in 7 Year Old Children. *Pediatrics* 132 (5), 1184-1193

Ketende, S. & Jones, E. (2011*) User Guide to Analysing MCS Data Using STATA.* Centre for Longitudinal Studies, Institute of Education.

Khandker, S., Koowal, G. & Samad, H. (2010) *Handbook on impact evaluation: quantitative methods and practises.* World bank, Washington.

Kiss, D. (2013) Impact of peer achievement and peer heterogeneity on own achievement growth. *Economics of Education Review* (37), 58-65

Kling, J. Liebman, J. & Katz. L (2007) Experimental analysis of neighbourhood effects. *Econometrica* 75(1), 83-119

Kristoffersen, J. & Smith, N. (2013) *Gender Differences in the Effects of Behavioral Problems on School Outcomes.* IZA discussion paper No. 7410

Kulick, J. & Kulik, C. (1992) Meta –analytic findings on grouping programs. *Gifted child Quarterly* 36(2), 73-77

Kunz, J. Page, M. & Solon, G. (2001) Are Single-Year Measures of Neighborhood Characteristics Useful Proxies for Children's Long-Run Neighborhood Environment? *Economic Letters* 79(2), 231–37

Kutnick, P., Sebba, J., Blatchford, P., Galton, M., & Thorp, J. (2005) *The effects of pupil grouping: literature review.* Department for Education and Skills.

Lavy, V. Paserman, D. & Schlosser, A. (2011) Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal* 122, 208-237

Lavy, V. Paserman, D. and Schlosser, A. (2011) Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal* 122, 208-237

Lavy, V. Silva, O. & Weinhardt, F. (2012) The Good, the Bad and the Average: Evidence of Ability Peer Effects in Schools. *Journal of Labor Economics* 30(2),367-414

Leadbeater, B. Kuperminc, G. & Hertzog, C. (1999) A Multivariate Model of Gender Differences in Adolescents' Internalizing and Externalizing Problems. *Developmental Psychology* 35(5), 1268-1282

Leis, J. Heron, J. Stuart, E. & Mendelson, T. (2014) Associations between maternal mental health and child emotional and behavioural problems: does prenatal mental health matter? *Journal of abnormal child psychology* 42(1):161-171

Leuven, E. (2014*) Help for psmatch2*. Retrieved 6th Jan 2015 from http://repec.org/bocode/p/psmatch2.html

Lewis, H., Hope, S.& Pearce, A. (2014) *Socioeconomic inequalities in parent-reported and teacher reported psychological well-being.* Retrieved 1st June 2015 from http://adc.bmj.com/content/early/2014/08/27/archdischild-2014-306288.full.pdf+html. Archives of Disease in Childhood

Lindahl, L. (2011) A comparison of family and neighbourhood effects on grades, test scores, educational attainment and income- evidence from Sweden. *The Journal of Economic Inequality* 9, 207-226

Ludwig, J. Kling, J. katz, L. sanbonmatsu, L. Liebman,J. Duncan,G. & Kessler, R. (2008) What can we learn about neighbourhood effects from the moving to opportunity experiment. *American Journal of Sociology* 114(1), 144-188

Lupton, R. (2003) *Neighbourhood effects can we measure them and does it matter?* CASE paper, 73. Centre for Analysis of Social Exclusion, London School of Economics and Political Science, London, UK.

Lupton, R. (2006) *How does place affect education? in Going Places: Neighbourhood, Ethnicity and Social Mobility*. London: Institute for Public Policy Research, 2006.

Lupton, R. and Kneale, D. (2010) *Are there neighbourhood effects on teenage parenthood in the UK, and does it matter for policy? A review of theory and evidence*. Centre for Analysis of Social Exclusion CASE paper 141

Machin, S. & McNally, S. (2008) The literacy hour. *Journal of Public Economics* 92, 1441-1462

Machin, S. & Silva, O. (2013) *School structure, school autonomy and the tail.* Centre for Economic Performance special paper no. 29.

Machin, S. & Vernoit, J. (2010) *CEP Briefing - A Note on Academy School Policy. The London School of Economics and Political Science,* Center of Economic Performance, London, UK.

Manley, D. & Van ham, M. (2010) *Neighbourhood effects housing tenure and individual employment outcomes.* IZA discussion paper no.5271

Manley, D., Van ham, M. & Doherty, J. (2011) *Social Mixing as a Cure for Negative Neighbourhood Effects: Evidence Based Policy or Urban Myth?* IZA Discussion Paper No. 5634

Marsh, H. & O'neill, R. (1984) Self-Description Questionnaire III: The Construct Validity of Multidimensional Self-Concept Ratings by Late Adolescents. *Journal of Educational Measurement* 21(2), 153-174

McClellan, E., & Kinsey, J. (1999). Children's social behavior in relation to participation in mixed-age or same-age classrooms. *Early Childhood Research and Practice*, 1(1), 1-19

McCulloch, A. & Joshi, H. (2001) Neighbourhood and family influences on the cognitive ability of children in the British national child development study. *Social Science and Medicine* 53(5),579-91.

McNeish, Di & Scott, S. (2014) Women and girls at risk: Evidence across the life course. DMSS retrieved 11th March 2015 from http://www.lankellychase.org.uk/assets/0000/2675/Women___Girls_at_Risk_-_Evidence_Review_040814.pdf

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy* 66 (4): 281–302.

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.

Mitchell, S. & Shepard, M. (2011) A comparative study of children's behaviour at home and at school. *British Journal of Educational Psychology* 36(3), 248-254

Mowbray, C., Lewandowski, L. Bybee, D.& Oysterman, D. (2005) Relationship Between Maternal Clinical Factors and Mother-Reported Child Problems. *Community Mental Health Journal* 41(6), 687-704

National Audit Office (2007) *The academies programme.* London.

National Audit Office (2010) *The academies programme.* London.

National centre for social Research (NatCen) (2008) *Child of the New Century Age 7 Survey Teacher Questionnaire – England.* Retrieved 9th February 2015 from http://www.esds.ac.uk/doc/6848/mrdoc/pdf/mcs4_teacher_england.pdf

Nicoletti, C. & Rabe, B. (2010) *Inequality in pupil's educational attainment: how much do family, sibling type and neighbourhood matter?* ISER Working Paper Series: 2010-26

Oakes, J. (1985) *Keeping track: How schools structure inequality.* New Haven: Yale University Press.

Oana, C. & Florent, S. (2012) Neighbourhood effects and unemployment outcomes: empirical evidence from French priority neighbourhoods. *Urban Public Economics Review* 17, 12-55

OECD (2014) *Are boys and Girls equally prepared for life?* Retrieved 18th March 2015 from http://www.oecd.org/pisa/pisaproducts/PIF-2014-gender-international-version.pdf

Office For Standards In Education (1998) *Setting in primary schools: A report from the Office of Her Majesty's Chief Inspector of Schools*. London: HMSO

Opdenakker, M. & Van Damme, J. (2001). Relationship between school composition and characteristics of school process and their effect on Mathematics achievement. *British Educational Research Journal* 27(4), 407–432

Oreopolous, P. (2003) The long run consequences of living in in a poor neighbourhood. *Quarterly Journal of Economics* 118(4), 1533-75

Overman, H. (2002) Neighbourhood effects in large and small neighbourhoods. *Urban Studies* 39(1), 117-130

Owens, A. (2010) Neighborhoods and Schools as Competing and Reinforcing Contexts for Educational Attainment. *Sociology of Education* 83(4), 287-311

Peters, M., Seeds, K., Goldstein, A.& Coleman, N. (2007) *Parental involvement in children's education 2007*. Department for children, schools and families research report. DCSF-RR034

Petts, R. (2009) *Fathers' religious involvement and early childhood behavior.* Fragile Families Working Paper 2009-22-FF

Platt, L. (2014) *Millennium Cohort Study Age 11 Survey Initial Findings.* London: Centre for Longitudinal Studies.

Plowden Report (1967) *Children and their primary schools*. London: HMSO

Pollard, A.& Triggs, P (2000) *What pupils say. Changing policy and practise in primary education.* Bloomsbury Academic

Popli, G. & Tsuchiya, A. (2014) *Sons and Daughters: Parental Beliefs and Child Behaviour (Evidence from the UK Millennium Cohort Study).* Working Papers 2014013. The University of Sheffield, Department of Economics.

Rasbash, J., Leckie, G., Pillinger, R. & Jenkins, J. (2010) Children's educational progress: partitioning family, school and area effects. *Journal of the Royal Statistical Society* 173 (3), 657–682

Reinherz, H., Paradis, A., Giaconia,R., Stashwick, C. & Fitzmaurice, G. (2003) Childhood and adolescent predictors of major depression in the transition to adulthood. *The American journal of psychiatry* 160(12), 2141-2147

Ritter, J. & Hancock, D.(2007) Exploring the relationship between certification sources, experience levels, and classroom management orientations of classroom teachers. *Teaching and Teacher education* 23(7), 1206-1216

Robertson, D. & Symons, J.(2003) Do peer groups matter? Peer groups versus schooling effects on Academic Attainment. *Economica* 70, 31-53

Rosenbaum P. & Rubin, D. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1), 41–55.

Rosenbaum, P. (1995) *Observational Studies.* Springer, New York

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Ryan, C. & Homel, J. (2014) *Educational outcomes: the impact of aspirations and the role of student background characteristics.* Research report no.65 LSAY

Sacerdote, B. (2001) Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics* 116 (2), 681-704.

Saleh, M. Ard, W & Jond, T. (2005) Effects of within-class ability grouping on social interaction, achievement and motivation *Instructional science* 33, 105-119

Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford,I. Taggart, B. Barreau, S. & Grabbe, Y. (2007) *Influences on Children's Development and Progress in Key Stage 2:*

Social/behavioural outcomes in Year 5. Department for Children, Schools and Families. Research Report DCSF-RR007

Sanbonmatsu, L., Kling, J. Duncan, G & brooks-Gunn, J. (2006) *Neighbourhoods and academic achievement: results from the moving to opportunity experiment.* The National Bureau of Economic Research working paper no. 11909

Schoon, I. (2006) Risk and Resilience: Adaptations in Changing Times. Cambridge: Cambridge University Press.

Scott, S., Doran, M. Beckett, C. Harry, S. & Cartwright, S. (2010) How is parenting style related to child antisocial behaviour? Preliminary findings from the Helping Children Achieve Study. Department for Education Research Report DFE-RR185a

SDQ info (2016) *Searchable database of SDQ publications.* Retrieved 18th October 2016 from http://www.sdqinfo.org/py/sdqinfo/f0.py

SDQ info (2014a) *What is the SDQ?* Retrieved from http://www.sdqinfo.com/a0.html 10th Jan 2015

SDQ info (2014b) *English (UK).* Retrieved 8th Jan 2015 from http://www.sdqinfo.com/py/sdqinfo/b3.py?language=Englishqz (UK)

Shavelson, R. & Bolus, R. (1981) Self-concept: The interplay of theory and methods. *Journal of educational psychology* 74(1), 3-17

Shin, Y. (2007) Peer Relationships, Social Behaviours, Academic Performance and Loneliness in Korean Primary School Children School. *Psychology International* 28(2), 220-236

Silva, O. (2006) *Faith Primary Schools: Better Schools or Better Pupils?* Centre for the Economics of Education

Slavin (1988) Synthesis of research on grouping in elementary and secondary schools. *Educational leadership.* 46(1),67-77

Solon, G. Page, M. & Duncan, G. (2000) Correlations between neighbouring children in their subsequent educational attainment. *The Review of Economics and Statistics* 82(3), 383-392

South, S. & Crowder, K. (1998) Leaving the 'Hood: Residential Mobility between Black, White, and Integrated Neighborhoods. *American Sociological Review 63* (1), 17-26.

Stock, J. Wright, J & Yogo, M.(2002) A survey of weak instruments and weak identification in generalizes methods of moments. *Journal of Business and Economic Statistics* 20, 518-529

Strand, S., Malmberg, L. & Hall, J. (2015). *English as an Additional Language (EAL) and Educational Achievement in England: an Analysis of the National Pupil Database*. London: EEF Retrieved from: http://educationendowmentfoundation.org.uk/uploads/pdf/EAL_and_educational_achievement2.pdf

Sukhnandan, L. & Lee, B. (1998) *Streaming, setting and grouping by ability: a review of the literature*. Slough: NFER.

Sulloway, F. (1996) *Born to rebel: Birth order, family dynamics, and creative lives*. New York, NY, US: Pantheon Books

TES (2013) *Streaming is out - now you can go with the flow*. Retrieved 10th Jan 2015 from https://www.tes.co.uk/article.aspx?storycode=6329982

Thomson, D. (2013). *Balancing Freedom and Fairness: Some Reflections on the Characteristics of Secondary Age Pupils Attending Academies.* FFT Research Paper 1

Van Ham M., Manley D., Bailey N., Simpson L. & Maclennan, D. (2012) *Neighbourhood Effects Research: New Perspectives.* Springer

Vignoles, A. & Meschi, E. (2010) *The Determinants of Non-Cognitive and Cognitive Schooling Outcomes*. Centre for the Economics of Education

Wang, L. (2014) All work and no play? The effects of ability sorting on students' non-school inputs, time use and grade anxiety. *Economics of Education Review* 44, 29-41

Weinhardt, A. (2013) *Neighbourhood quality and student performance*. IZA discussion paper 7139

Whitburn, J. (2001) Effective classroom organisation in primary schools: mathematics. *Oxford Review of Education*, 27 (3), 411-428

Wilson, D. & Piebalga, A. (2008) *Accurate performance measure but meaningless ranking exercise? An analysis of the English school league tables*. Centre for market and public organisation working paper 07/176

Wilson, J. (2011). *Are England's Academies More Inclusive or More 'Exclusive'? The Impact of Institutional Change on the Pupil Profile of Schools.* CEE Discussion Paper 125

Wilson, W (1987) *The truly disadvantaged: the inner city, the underclass and public policy. Chicago.* University of Chicago Press

Worth, J. (2014) *Analysis of academy school performance in GCSEs 2014*. NFER. Final report

Xu, L., Gauthier, A. & Strohschien, L. (2009) Why are some Children Left Out? Factors Barring Canadian Children from Participating in Extracurricular Activities. *Canadian Studies in Population* 36(3-4)325-345